5-2020

Visualization of Structural Changes in Thermochemically Pretreated Plant Cell Wall Polymers

Riddhi Shah
University of Tennessee, rshah10@vols.utk.edu

Recommended Citation

Shah, Riddhi, "Visualization of Structural Changes in Thermochemically Pretreated Plant Cell Wall Polymers. " PhD diss., University of Tennessee, 2020.
https://trace.tennessee.edu/utk_graddiss/5839

To the Graduate Council:

I am submitting herewith a dissertation written by Riddhi Shah entitled "Visualization of Structural Changes in Thermochemically Pretreated Plant Cell Wall Polymers." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Energy Science and Engineering.

Hugh O'Neill, Brian Davison, Major Professor

We have read this dissertation and recommend its acceptance:

Arthur Ragauskas, Nicole Labbe, Mark Dadmun

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Visualization of Structural Changes in Thermochemically Pretreated Plant Cell Wall Polymers

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Riddhi Shalibhadra Shah
May 2020

Acknowledgments

I am grateful to the following people without whose help I would not have been able to successfully complete my doctoral degree program.

I would like to start by thanking my department, Bredesen center for Interdisciplinary Science and Research former director Dr. Lee Riedinger for accepting me as a doctoral student in this program. I would also like to thank the current director Dr. Suresh Babu and the most caring, diligent and affectionate administrative staff including Wanda Davis, Jessica Garner and Allie Burns of the Bredesen center.

Words are not enough to express the respect and gratitude I have for my Ph.D. mentors; Dr. Hugh O’Neill and Dr. Brian H. Davison. I joined them as a research associate in 2013 and I fancied a dream of becoming a scientist as I observed how they addressed and solved problems. I became their Ph.D. student from 2015 and through the course of this training period, they taught me the most important lessons of life like working hard, never giving up, thinking critically and independently and lastly believing in myself. They were always available and spend time with me either discussing results, reading my publications or proposals or listening to mock presentations before conferences as many times as I wanted to practice. As my degree program comes to an end, I will miss their guidance the most but they have left a big influence on me and they shall forever have a special place in my heart. I aspire to be a good scientist and just as good a person as my mentors.

I want to thank the members of my dissertation committee Prof. Arthur J. Ragauskas, Prof. Nicole (Nikki) Labbe and Prof. Mark Dadmun for their valuable time and input to my growth and dissertation. I want to thank Prof. Ragauskas for teaching a class on biorenewable polymers. This class was useful in getting an understanding of the chemistry of plant cell wall polymers. I want to thank Dr. Labbe for the knowledge she shared with me related to lignin and plant cell wall. I want to thank Dr. Dadmun for pointing out chapters in a book that were a starting point for me to learn small-angle scattering. I want to extend a special thank you to Prof. Claudia Rawn who taught a class on X-ray diffraction. The concepts taught by her in that class were critical for me to learn small-angle scattering.

I want to thank the Department of Energy's Biological and Environmental Research (BER) program that funded the Biofuels- scientific focus Area led by Dr. Brian H. Davison. The entire team of this project has in some way helped during my Ph.D. In particular, I want to thank Dr. Barbara R. Evans who gave so many useful insights to my work and has been so supportive, motivating and enthusiastic through the course of my Ph.D. I also want to thank Dr. Sai Venkatesh Pingali who has been so patient with me and helped me with all the neutron scattering studies. I want to thank Dr. Jong Keum from the Center for Nanophase Material Science (CNMS) at ORNL for teaching me few basic but most important concepts of small angle scattering and diffraction. I also want to thank him for his unconditional user support whenever I used X-ray instruments in his lab. I want to thank Zachary Gosser and Dr. Adam Rondinone for letting me use the CNMS lab equipments. I want to thank Miguel Rodriguez Jr. for helping me with sugar analysis in the composites and feedstocks. (Also, for sharing the best coffee and coffee creamer)

I want to thank the user support staff at the High Flux Isotope Reactor (HFIR) and Spallation Neutron Source (SNS). Specially I want to thank the instrument scientists of beam line CG3/BioSANS Dr. Sai Venkatesh Pingali, BL6/EQ-SANS Dr. William Heller and Dr. Changwoo Do, and USANS Dr. Ken Littrell for help with data collection and reduction. I want to thank Douglas Armitage for making it easier for me to use the McHugh pressure cell. I also want to thank sample management personnel Kristin Nevius and radiological control technician Phong Nguyen for their help during beam time experiments. I want to thank Jaimie Werner and Rhonda Moody for their help during neutron experiments and during sample preparation. Most of the sample preparation and characterization work was done at the Center for Structural and Molecular Biology (CSMB) labs and I want to thank Dr. Kevin Weiss for all his suggestions and inputs to my work.

I want to thank all my friends and in particular Dr. Daisuke Sawada, Dr. Nayomi Plaza, Dr. Zhi Yang, Dr. Margarita Fomina, and Dr. Pan Chen who were either in their post-doctoral training or doctoral training during my time as a graduate student at ORNL. Discussions about the field with them were valuable and broadened my horizon of thinking. I want to thank Dr. Mini Malhotra, Dr. Amandeep Sangha, and Dr. Venugopal Vandavasi for always lifting my spirits and motivating me.

It is often said you do good you get good and otherwise. This understanding fails in one instance as I have done nothing so good to deserve the parents such as mine. I am fortunate and blessed to have parents that are so loving, caring and supportive. I also want to thank my brother, sister-in-law and my almost two-year-old nephew for their constant support and love.

Abstract

Second-generation biofuels made from lignocellulosic biomass hold immense potential in serving as an alternative source of energy. Due to the rigid cell wall structure, the biomass has to be pretreated with chemicals, often at high temperature and pressure, to breakdown the cell wall structure and increase cellulose accessibility to hydrolytic enzymes. Through years of research, a great amount has been learned about the structural rearrangements that occur after pretreatment and have resulted in proposed reasons for recalcitrance to enzyme hydrolysis that occurs even after pretreatment. But why the structural re-arrangement took place the way it did during pretreatment is largely unknown. This is because our current understanding of how polymers are interacting and influencing each other's structures during pretreatment is lacking. The overall aim of the work is to understand how plant cell wall polymers interact in lignocellulose and the changes that occur in these interactions during hot water or dilute acid pretreatment. Our approach involves studying polymer-polymer interactions using model systems and conducting comparative studies in natural variants of poplar and switchgrass. Using x-rays and neutrons based scattering techniques, we studied the impact different kinds of hemicelluloses have on the hierarchical structure of cellulose before and after pretreatment. We also studied intact biomass and used a comparative approach to study the sugar release differences before and after pretreatment for two poplar woods with different lignin content. Lastly, we also explored the interactions of pectin with lignin to determine if pectin can influence the lignin aggregate formation. Overall, information on this fundamental knowledge of polymer interactions can help design milder pretreatments or plants that are more susceptible to deconstruction, which would decrease the cost of thermochemical pretreatment.

Table of Contents

List of Tables

List of Figures

Acronyms

1D	One-Dimensional
2D	Two-Dimensional
5-HMF	5- Hydroxyl Methyl Furfural
AFEX	Ammonia Fiber Explosion
AFM	Atomic Force Microscopy
ARP	Ammonia Recycle Percolation
ATR	Attenuated Total Reflectance
CARS	Coherent Anti-strokes Raman Scattering
CBM	Carbohydrate Binding Module
COMT	Catechol O-Methyl Transferase
D	Deuterium
DAP	Dilute Acid Pretreatment
DB	Direct Blue
DHP	Dehydrogenation Polymer
DO	Direct Orange
DP	Degree of Polymerization
EAP	Enhanced Angle Pressure
EISA	Energy Independence and Security Act
EPA	Environmental Protection Agency
FWHM	Full Width Half Maximum
G	Guaiacyl
GAUT4	Galacturonosyl Transferase 4
GHGs	Green House Gases
GM	Glucomannan
GPC	Gel Permeation Column
GWAS	Genome Wide Association Study
H	Hydrogenated/Protiated
HFIR	High Flux Isotope Reactor
HG	Homogalacturonan
HPLC	High Performance Liquid Chromatography
HRP	Horse Radish Peroxidase
IR	InfraRed
Kd	Knock down
LCCs	Lignin Carbohydrate Complexes
LHW	Liquid Hot Water
Mw	Molecular weight
NIST	National Institute of Standards and Technology
NMR	Nuclear Magnetic Resonance
ORNL	Oak Ridge National Laboratory
PCW	Plant Cell Wall
RFS	Renewable Fuel Standards
R_g	Radius of Gyration
RGI	Rhamnogalacturonan I
RI	Refractive Index

S	Syringyl
SAA	Soaking in Aqueous Ammonia
SANS	Small Angle Neutron Scattering
SAXS	Small Angle X-ray Scattering
SEM	Scanning Electron Microscopy
SFG	Sum Frequency Generation
SLD	Scattering Length Density
SNS	Spallation Neutron Source
TEM	Transmission Electron Microscopy
Tg	Glass Transition Temperature
USANS	Ultra-Small Angle Neutron Scattering
UV	Ultra Violet
WAXS	Wide Angle X-ray Scattering
XG	Xyloglucan
XRD	X-ray Diffraction

Chapter 1). Introduction

1.1). Need for biofuel

1.1.1). Consumption of petroleum

Energy from sources such as petroleum, natural gas, coal, nuclear and renewables are consumed by the transportation sectors in the form of fuel and the industrial, commercial and residential sector in the form of electricity.[1] In 2018, 36% of the total energy consumed in the United States was from petroleum making it the largest primary energy source of the U.S.[2] Petroleum comprises of crude oil, a naturally occurring liquid that is found in geological formations beneath the earth surface.[3] Crude oil consists of hydrocarbons, organic compounds and small amount of metals and it is refined to form many petroleum products such as gasoline, distilled fuel oil such as diesel and heating oil, jet fuel, coke, lubricants, kerosene.[4] Petroleum products are used for deriving other value-added products such as plastics, adhesives, synthetic rubbers, polyurethane, solvents, etc.[5] In 2018, 20.50 million barrels per day (average) petroleum was consumed in the U.S., of which two-thirds of finished petroleum products were used for transportation purposes.[4] Figure 1.1 shows a further breakdown of the most used petroleum product.

1.1.2). Greenhouse gases and effect on the environment

Due to burning of the fossil fuels like coal and petroleum, there is a rapid increase and a buildup in the content of the greenhouse gases in the earth's atmosphere.[7] The greenhouse gases (GHG) such as water vapor, ozone, carbon dioxide (CO_2) and other greenhouse gases like methane (CH_4), nitrous oxide (N_2O) are naturally present in the earth's atmosphere and they entrap the sun's heat causing the greenhouse effect.[8]

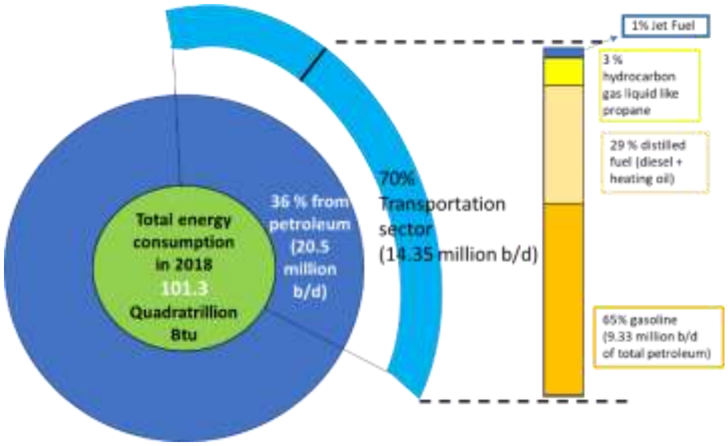


Figure 1.1: A breakdown of the total energy consumed (unit of energy in British Thermal Units)to show the usage in the transportation sector and a further breakdown to show gasoline as the leading product consumed in the transportation sector. (Values obtained from Monthly Energy Review, Energy Information Administration)[1, 4, 6]

Due to this naturally occurring greenhouse effect suitable temperature for life on earth is maintained. However, the content of these gases in earth's atmosphere has increased by 37% from 1990 to 2015 which has led to an increase in temperature by almost 0.5˚C.[9-10] Burning of transportation fuel accounts for 29% of total U.S. greenhouse gas emissions.[11] It is important to note that the common mode of transportation is roadway in the U.S due to which the GHGs emissions from burning petroleum account for most of U.S. GHG emissions. In fact, 60% of the total emissions are from light-duty vehicles.[12] Additionally, the recent trend in the consumption of coal, the other leading source of primary energy in the U.S has seen a downward fall but consumption of petroleum is seeing an increase. This indicates that the amount of CO_2 from burning petroleum is also increasing in the earth's atmosphere (Figure 1.2).

Today the atmospheric content of carbon dioxide is higher than it has ever been in the past 400,000 years accounting for a total of 400 parts per million (ppm). This is projected to lead to a drastic increase of 4˚C in earth's temperature by 2100.[14] The effects of this increase in temperature would result into increasing aridity and drought in some area, major floods in other areas and a sea-level increase of 0.5 to 1 meter along with increase in the acidic levels of ocean water. All these situations will negatively impact agricultural production, result in loss of habitat for all land-living species and death of sea-creature due to increased acidity of water.[8, 14]

As the consequences of climate change become well-known, the Environmental Protective Agency (EPA) has introduced policies and programs to curb its consequences. To name a few these include setting GHGs emissions standard for cars and trucks, working with freight transportation sector to improve supply chain efficiencies, buying government-owned low emission vehicles at federal agencies and taking steps to set emission standards for aircraft.[15] The most significant policy has been the Renewable Fuel Standard (RFS) program due to which the nation's renewable fuels sector

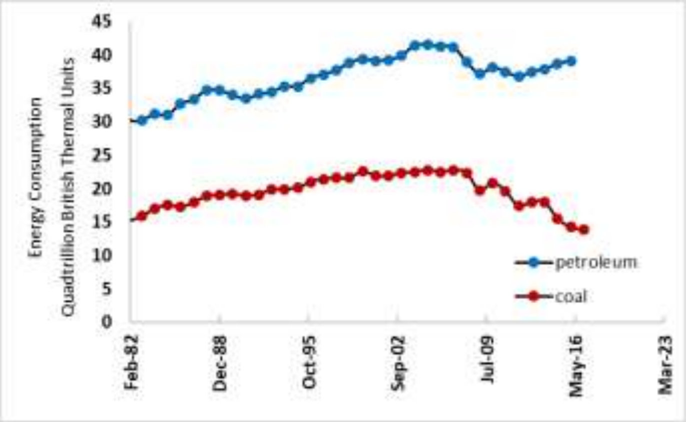


Figure 1.2: Graph showing a downfall in consumption of coal (red dots) and while increase in the consumption of petroleum (blue dots) as a primary energy source[13]

has expanded. According to this program the fuel for vehicles must include ethanol-based fuels made from plants also called biofuels.[16] Biofuels offer several benefits over fuel made from petroleum.[17] Plants can be grown, harvested and converted to fuel over and over which makes biofuels a renewable source of energy. The production of the fuel from wide variety of plant materials and waste will be domestic which would result in decrease of the energy dependence of the nation on other countries. Plants utilize CO_2 during growth and may increase stores of soil organic carbon resulting in decrease of GHG emissions relative to petroleum-derived fuels. Additionally, vehicle tailpipe emissions of many air pollutants harmful to human health may be lower with biofuels. Lastly, biofuels production and use have provided more jobs in sectors ranging from agriculture, transportation and in a whole new industrial sector for converting and refining plant-based fuel.[17]

1.2). Biofuels

1.2.1). First generation biofuel

Any plant rich in the starting material for ethanol production namely, fermentable sugars, would serve as an excellent feedstock.[18] The 1st generation biofuel in US are mainly derived from corn which is rich in starch, a polymer made of α (1→4) linked glucose. Historically, ethanol from corn was proposed to be used as a transportation fuel but due to lower cost and higher energy output gasoline was used instead.[19] The use of corn for making ethanol gained traction in the US during late 1990s when methyl tertiary butyl ether, an oxygenate additive to gasoline was responsible for contaminating ground water.[20] Around this time a blend of fuel containing 10% of ethanol to gasoline (E10) was replacing 100% gasoline commercially. Since ethanol has a higher-octane number (106–110) than gasoline (91–96), the blend was able to withstand higher compression and the fuel would ignite near the ignition spark in the compression chamber.[21] Additionally, in 2002 flex-fuel vehicles that would run with up to 85% ethanol were commercially being made by Ford which further increased ethanol production from corn. In 2005 as an amendment to the Clean Air Act, the renewable fuel standards were established and a limit was set to include a minimum volume of renewable fuel such as ethanol to gasoline.[22] In the same year, the US became the world's largest producer of ethanol fuel surpassing Brazil, the only other country which was significantly invested in the production of ethanol albeit from sugar-cane.[23] However, as corn is used as fodder for animals and for cooking purposes by humans it was soon realized that corn-based biofuel might negatively impact the food chain, lead to over consumption of water and a loss of agricultural land. Other sugar-based feedstocks had same problems as corn for deriving 1st generation biofuels and these included sugarcane (usually grown in tropical climate), sugar beet, molasses obtained from sugar cane and sugar beet, or unconsumed fruits.[18] Around 2006 the prices for gasoline had skyrocketed and under the Energy Independence and Security Act of 2007 a further increase in the production of biofuel was mandated. This policy delineated the quantities of biofuels that would be derived from corn and those that would be derived from lignocellulosic biomass that is not traditionally consumed by humans.[16] As per the EISA act of 2007, a target of producing 36 billion gallons of biofuel by 2022 was set. As shown in Figure 1.3, the amount of ethanol from corn was capped at about 16 billion gallons and by 2022, around 20 billion gallons of ethanol were supposed to be made from cellulosic sources.

In 2018, about 14.4 billion gallons of fuel ethanol were consumed in the United States and 95% of the ethanol was derived from corn.[24]

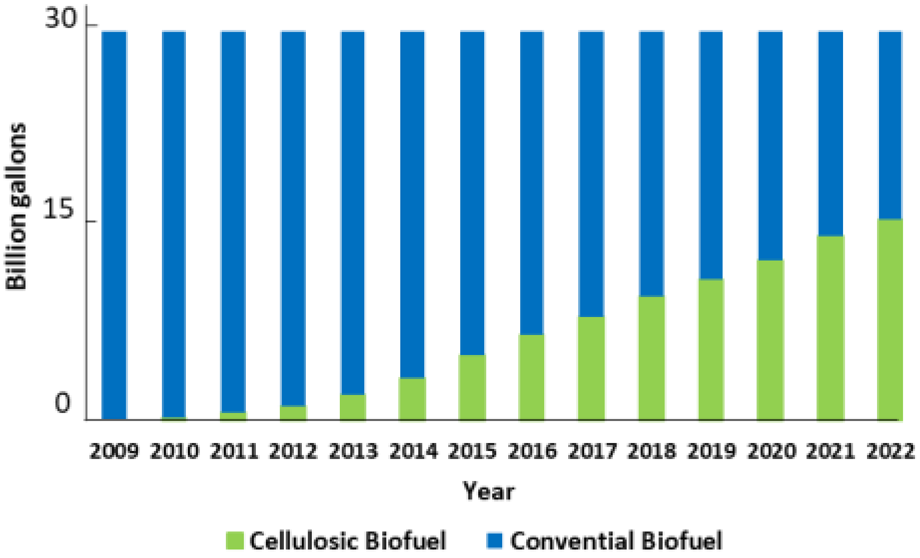


Figure 1.3: The target of cellulosic based and conventional biofuel set by according to the EISA of 2007 from 2009 to 2022[16]

While the production of cellulosic ethanol did not meet its target for 2017 and would possibly not meet its target for 2022,[25] it is worth noting that because of the production of corn-based fuel and its addition to the E10 blend, gasoline consumption decreased. Corn-based ethanol does reduce GHGs emissions by 20% as compared to petroleum.[11] However, in addition to being a crop for food consumption, they are also not as cost-effective as gasoline. Additionally, the bio-refineries that make this biofuel are run with coal. There is a need to replace corn-based ethanol by other advanced biofuel sources such as cellulosic ethanol or from ethanol from algal sources. Both these advanced biofuels would reduce GHGs by 60% and 50% as compared to petroleum.[11]

1.2.2). Advanced biofuel

While initially there was research conducted to turn corn to ethanol since the early 2000s research is mainly focused on conversion of lignocellulosic biomass to biofuel. Pilot-scale plants and biorefineries for cellulosic ethanol are growing.

They use lignin instead of coal to provide energy to run the biofuel plant.[19] However, a huge gap in translating the research to commercial market exists. One of the prime reasons is that 2nd generation biofuels (feedstock was sorghum) are still more expensive than gasoline (cost ranges from >50% to almost comparable under the most optimistic scenario).[26] It is important to note that the advantages and cost of biofuels over fossil fuels are largely dependent on the process by which the biofuel is produced. Several factors such as energy inputs to operate a bio-refinery, rate at which plant biomass is grown, amount of biomass made, conversion to biofuel efficiency, simultaneous development pollution control technologies like carbon sequestration, and direct and indirect land-use have implications on the cost and the advantage biofuels bring to a sustainable environment.[17, 27-28] In order to understand these bottlenecks that drive up the costs, we need to understand the process of conversion of lignocellulosic biomass to biofuel

The conversion of plant biomass to 2nd generation cellulose-based biofuel at a bio-refinery[29] (Figure 1.4) can be broken down into following steps:

1). Selection growth, harvest, storage and transportation of feedstock

For making 2nd generation cellulose-based biofuel, feedstocks comprise of lignocellulosic materials derived from agricultural residues (few examples include corn stover, sugarcane bagasse, rice, and wheat straw)[30] and energy crops (few examples include switchgrass, poplar, miscanthus, willow).[31-32] The common and preferable characteristics for any feedstock to be considered for biofuel production are they should be non-food crops, easy to grow with less energy input, give high yield in less amount of time, and can withstand unfavorable conditions such as drought, salt stress, high temperatures, and flooding. Also, the selection of the bioenergy crop is largely based on the type of geographical location and environmental conditions.[33]

2). Pretreatment of the plant biomass:

Due to the rigid cell wall structure of plant biomass, it has to be pretreated either by physical, chemical or biologically methods. More often than not the feedstocks are cut to a certain size by milling and then treated with solvents. The function of the pretreatment process is to open the lignin-hemicellulose matrix and disrupt the cellulose ultrastructure making the lignocellulosic biomass more accessible for the hydrolytic enzymes Pretreating biomass with solvents at high temperature conditions results in separating the polymer cellulose and rest of the biomass components.[34-35]

3). Enzymatic hydrolysis of cellulose to cellobiose and glucose

The extracted cellulose is treated with hydrolytic enzymes to break cellulose down to glucose. Enzymatic hydrolysis is a heterogeneous reaction that involves multiple steps.

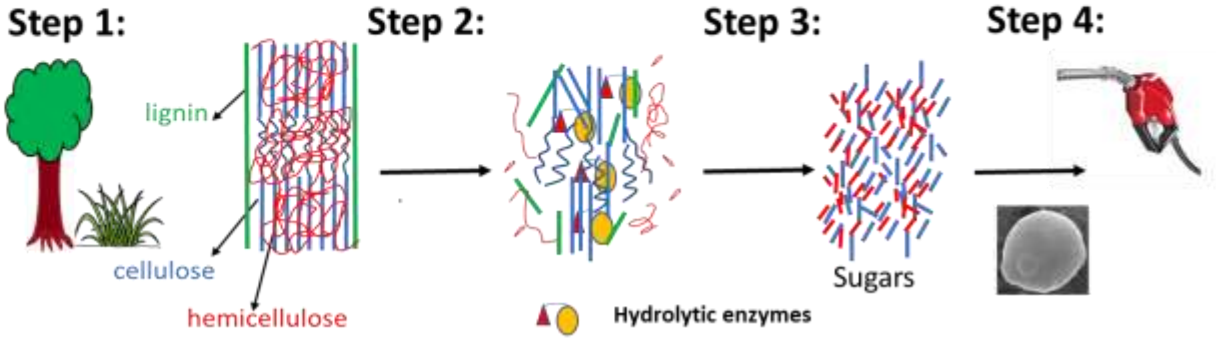


Figure 1.4: The schematic shows the four steps -1). Selection of feedstock 2). Pretreatment of feedstock or lignocellulosic biomass 3). Enzymatic hydrolysis of cellulose 4). Fermentation of glucose for conversion to ethanol

Glycosidic hydrolases (GHs) are a class of hydrolytic enzymes that responsible for cleaving the glycosidic bond found in polysaccharides such as cellulose. They contain the endoglucanase and exoglucanase that work together on the insoluble cellulose at the solid-liquid interface. Due to their action cellobiose or other short oligosaccharides are formed which enter the liquid medium. In the next step, these soluble sugars are further catalytically acted upon by β- glucosidase and cleaved to glucose.[36]

4). Fermentation of glucose to alcohols such as ethanol

The last step for conversion includes fermenting glucose to ethanol. The solution rich in sugars formed from the previous step is recovered and acted upon by micro-organisms (to name a few – *Saccharomyces*, *E. coli*, and *Zymomonas*). These microbes utilize the sugar and release ethanol. The ethanol is then separated and purified by distillation.[37]

While each step of the process to produce biofuel have problems that need to be addressed, the step 2 contribute to about 30% of the total cost of production.[35] Commonly used industrial thermochemical pretreatment approaches use hot water or dilute acid for the breakdown of the cell wall structure. These solvents also hydrolyze the matrix polysaccharides to oligomeric or monomeric sugars by acting on the glycosidic linkage. In these pretreatment approaches, water or dilute acid (0.2 to 2.5% w/w) is added to the lignocellulosic biomass which is then heated to 120 - 200 ˚C with constant mixing in a closed reactor.[38] However, the need for high pressure and temperature increase the cost of the process and is seen as a major disadvantage for the production of second-generation biofuels using this pretreatment approach.[39]

For optimizing the water or dilute acid pretreatment, it is critical to determine the structure of the cell wall and type of changes that occur to the cell wall as a result of pretreatment. In **Chapter 2**, a detailed description of the structure of individual cell wall polymers that make the plant cell wall, the proposed structural models of the primary and secondary cell wall, reasons for biomass recalcitrance to enzymatic hydrolysis, and pretreatments methods have been mentioned. In very simple terms, the cellulose is embedded in a matrix of polymers such as lignin and hemicellulose and post hot water or dilute acid pretreatment the cellulose structure collapses while lignin forms aggregates and hemicellulose partially get solubilized during pretreatment.[40-47] Although structural changes that occur in the plant cell wall due to pretreatment has been studied,[42, 48] less is known about the polymer interactions that lead to the assembly of the plant cell wall and those that drive these changes during pretreatment.

1.3). Objective

Hence, the focus of this work was to understand how plant cell wall polymers interact and assemble to make the complex plant cell wall structure and the changes that occur in these interactions during hot water or dilute acid pretreatment, eventually resulting in the decrease in recalcitrance.

We studied polymer-polymer interactions and the changes that result in them and during pretreatment using A). model systems and B). by conduct comparative studies in natural variants of poplar. These approaches were chosen due to the inherent complexity of the cell wall. The plant cell wall composition is highly variable not only from one plant to another but between different tissues of the same plant. The presence of the many different polymers with the associated presence of several functional groups makes the plant cell wall complex and heterogeneous. **Chapter 2** surveys plant cell wall structure and pretreatment. As a result of the heterogeneity it has been difficult to study polymer-polymer interaction, yet the development of the different analytical methods and techniques have enabled characterization of the molecular level interactions and chemical features both of individual cell wall polymers and in the cell wall. In this thesis, the lignocellulosic biomass was studied with scattering techniques such as small-angle neutron scattering (SANS), small and wide-angle X-ray scattering. **Chapter 3** describes the fundamentals of scattering for studying and analyzing the structure of lignocellulosic biomass and its components. The use of certain fitting approaches that were used for the scattering data is also mentioned. Specifically, coupling scattering with other chemical or physical characterization methods such as quantitative saccharification, infrared spectroscopy, cellulose accessibility measurements we studied hemicellulose-cellulose interactions (**Chapter 4**), lignin-cellulose interactions (**Chapter 5)** and lignin-pectin interactions (**Chapter 6)**.

In **Chapter 4**, we report the results of our study in which we designed a controlled model system and investigated the influence a particular polymer has on the structure of the other polymer. Especially, changes in cellulose structure in the presence of different kinds of hemicellulose were important to study to make better biofuels as cellulose is secreted into a matrix containing hemicellulose and also cellulose is the polymer that is eventually converted to biofuel. Hence, cell wall mimicking hemicellulose-cellulose composites were made and subjected to dilute acid pretreatment to study the influence of different hemicelluloses on the structure of cellulose both before and after pretreatment. The composites were prepared by synthesizing bacterial cellulose from *Acetobacter xylinus* in the presence of glucomannan or xyloglucan dissolved in the growth media. SANS was used to study changes in cellulose macrofibril size, surface, and internal morphology of macrofibril and arrangement of macrofibrils within the network like structure in untreated and dilute acid pretreated (DAP) cellulose and composites. The SANS results are complemented by the changes in crystallinity, crystallite size and peak position as detected by X-ray diffraction and sugar compositional analysis. Our published results point to a tight interaction of xyloglucan with microfibrils while glucomannan only interacts at the macrofibril surfaces.[49]

To understand cellulose and lignin interaction (**Chapter 5),** we looked at the plant cell wall as a whole. We wanted to gain insights into how the decrease of one polymer component influences changes in the other components eventually leading to changes in sugar yield upon pretreatment followed by enzymatic hydrolysis. As previously mentioned, the plant cell wall is complex and studying polymer interactions in the plant cell wall, and concluding can be a difficult task. But interdependency among the polymers suggests that changes in one polymer can lead to changes in the amount, structure, or other properties of other polymers in the cell wall. A comparative approach was taken to study structural differences in the cellulose due to lignin content of two naturally occurring poplar, which were grown under the same conditions, harvested at the same time, but showed differences in sugar release in the native and hot water pretreated state. These two kinds of poplar (kind of hardwood) were a part of the genome-wide association study (GWAS)[50,] and we found that the poplar with a lower content of lignin had a different distribution of lignin and arrangement of cellulose microfibrils than high lignin contenting poplar. Hot water pretreatments at three different time and temperature conditions were conducted, and simultaneously the changes occurring in the poplar were studied with SANS. Post pretreatment, no differences were found in the way lignin re-distributed, and the coalesced cellulose microfibrillar cross-section size between the two poplar samples for the same condition of pretreatment. Additionally, no difference in cellulose accessibility, the formation of inhibitors, or acetic acid after hot water pretreatment were observed. However, with X-ray diffraction (XRD) changes in the cellulose ultrastructure such as crystallite size, crystallinity and crystallinity between the two poplars were observed for the most severe pretreatment which could be the reason for the previously reported observed difference in glucan release.

In **Chapter 6**, we investigated pectic polysaccharide interaction with lignin. We conducted hot water pretreatment on wildtype and galacturonosyltransferase 4 (GAUT4) knockdown switchgrass to study differences in lignin aggregate size and used model systems to check for the possible formation of lignin-carbohydrate complexes. A recent study used monoclonal antibodies to track the location of galacturonic acid-containing homogalacturonan pectin.[51] Interestingly, homogalacturonan was found to be in the middle lamella and cell wall corner region as lignin. Sugar release from switchgrass mutant (GAUT4-knockdown) in which the gene responsible for making the galacturonic acid was suppressed resulted in an increase after hot water pretreatment compared to the wildtype switchgrass.[51] These results encouraged us to look for the lignin-carbohydrate complex formation between lignin and pectin and study its influence on the lignin aggregate formation using SANS. Further, to investigate the nature of interaction (physical or chemical), we made synthetic lignin from monolignols in the presence of homogalacturonan. We investigated the lignin-homogalacturonan composite with small-angle X-ray scattering (SAXS) and infra-red (IR) spectroscopy.

In summary (**Chapter 7**), by using plant cell wall polymers for making model materials and combining results with plant studies this work will provide valuable insights into the interactions between the plant cell wall polymers, which can be useful for future development of biofuels and bioproducts.

1.4). References

1. Primary energy consumption. https://www.eia.gov/totalenergy/data/monthly/pdf/flow/css_2018_energy.pdf (accessed April 2019).
2. In the United States, most petroleum is consumed in transportation. https://www.eia.gov/todayinenergy/detail.php?id=40752#.
3. Wolfson, R., *Energy, environment, and climate*. W.W. Norton & Company: New York, 2012.
4. Oil: crude and petrolum products explained. https://www.eia.gov/energyexplained/oil-and-petroleum-products/in-depth.php.
5. Alfke, G.; Irion, W. W.; Neuwirth, O. S., Oil Refining. *Ullmann's Encyclopedia of Industrial Chemistry* **2007**.
6. Energy use for transportation. https://www.eia.gov/energyexplained/use-of-energy/transportation.php.
7. Oreskes, N., The Scientific Consensus on Climate Change. *Science* **2004,** *306* (5702), 1686.
8. Jain, P. C., Greenhouse effect and climate change: scientific basis and overview. *Renewable Energy* **1993,** *3* (4), 403-420.
9. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2015*; 2016.
10. Meehl, G. A.; Stocker, T. F.; Collins, W. D.; Friedlingstein, P.; Gaye, T.; Gregory, J. M.; Kitoh, A.; Knutti, R.; Murphy, J. M.; Noda, A., Global climate projections. **2007**.
11. Fast Facts on Transportation Greenhouse Gas Emissions. https://www.epa.gov/greenvehicles/fast-facts-transportation-greenhouse-gas-emissions.
12. Barnett, J. R.; Bonham, V. A., Cellulose microfibril angle in the cell wall of wood fibres. *Biological Reviews* **2004,** *79* (2), 461-472.
13. Petroleum and coal consumption trends. https://www.eia.gov/opendata/.
14. *Turn Down the Heat: Confronting the New Climate Normal*. The World Bank.: Washington D.C., 2014.
15. Carbon Pollution from Transportation. https://www.epa.gov/transportation-air-pollution-and-climate-change/carbon-pollution-transportation.
16. Overview for Renewable Fuel Standard. https://www.epa.gov/renewable-fuel-standard-program/overview-renewable-fuel-standard.
17. Hill, J.; Polasky, S.; Nelson, E.; Tilman, D.; Huo, H.; Ludwig, L.; Neumann, J.; Zheng, H.; Bonta, D., Climate change and health costs of air emissions from biofuels and gasoline. *Proceedings of the National Academy of Sciences* **2009,** *106* (6), 2077.
18. Naik, S. N.; Goud, V. V.; Rout, P. K.; Dalai, A. K., Production of first and second generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews* **2010,** *14* (2), 578-597.
19. Pandey, A.; Negi, S.; Binod, P.; Larroche, C., *Pretreatment of Biomass: Processes and Technologies*. Elsevier Science: 2018.
20. Hartley, W. R.; Englande, A. J.; Harrington, D. J., Health risk assessment of groundwater contaminated with methyl tertiary butyl ether (MTBE). *Water Science and Technology* **1999,** *39* (10), 305-310.
21. Alternative fuel data center. https://afdc.energy.gov/fuels/ethanol.html.

22. McCarthy, J. E.; Copeland, C.; Parker, L.; Schierow, L.-J. In *Clean Air Act: A summary of the act and its major requirements*, 2007; Citeseer.

23. Industry Statistics. https://web.archive.org/web/20080408091334/http://www.ethanolrfa.org/industry/statistics/#E.

24. Today in energy. **2018**.

25. Proposed Volume Standards for 2019, and the Biomass-Based Diesel Volume for 2020. https://www.epa.gov/renewable-fuel-standard-program/proposed-volume-standards-2019-and-biomass-based-diesel-volume-2020.

26. van Rijn, R.; Nieves, I. U.; Shanmugam, K. T.; Ingram, L. O.; Vermerris, W., Techno-Economic Evaluation of Cellulosic Ethanol Production Based on Pilot Biorefinery Data: a Case Study of Sweet Sorghum Bagasse Processed via L+SScF. *BioEnergy Research* **2018,** *11* (2), 414-425.

27. Hill, J.; Nelson, E.; Tilman, D.; Polasky, S.; Tiffany, D., Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences* **2006,** *103* (30), 11206.

28. Ramos, J.-L.; Valdivia, M.; García-Lorente, F.; Segura, A., Benefits and perspectives on the use of biofuels. *Microb Biotechnol* **2016,** *9* (4), 436-440.

29. Brethauer, S.; Studer, M. H., Biochemical Conversion Processes of Lignocellulosic Biomass to Fuels and Chemicals - A Review. (0009-4293 (Print)).

30. Kadam, K. L.; McMillan, J. D., Availability of corn stover as a sustainable feedstock for bioethanol production. *Bioresource Technology* **2003,** *88* (1), 17-25.

31. Ye, X.; Busov, V.; Zhao, N.; Meilan, R.; McDonnell, L. M.; Coleman, H. D.; Mansfield, S. D.; Chen, F.; Li, Y.; Cheng, Z.-M., Transgenic Populus Trees for Forest Products, Bioenergy, and Functional Genomics. *Critical Reviews in Plant Sciences* **2011,** *30* (5), 415-434.

32. Keshwani, D. R.; Cheng, J. J., Switchgrass for bioethanol and other value-added applications: A review. *Bioresource Technology* **2009,** *100* (4), 1515-1523.

33. Phitsuwan, P.; Sakka, K.; Ratanakhanokchai, K., Improvement of lignocellulosic biomass in planta: A review of feedstocks, biomass recalcitrance, and strategic manipulation of ideal plants designed for ethanol production and processability. *Biomass and Bioenergy* **2013,** *58*, 390-405.

34. Lynd, L. R.; Liang, X.; Biddy, M. J.; Allee, A.; Cai, H.; Foust, T.; Himmel, M. E.; Laser, M. S.; Wang, M.; Wyman, C. E., Cellulosic ethanol: status and innovation. **2017,** (1879-0429 (Electronic)).

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

36. Yang, B.; Dai, Z.; Ding, S.-Y.; Wyman, C. E., Enzymatic hydrolysis of cellulosic biomass. *Biofuels* **2011,** *2* (4), 421-449.

37. Tomas-Pejo, E.; Oliva, J. M.; Ballesteros, M., Realistic approach for full-scale bioethanol production from lignocellulose: a review. **2008**.

38. Agbor, V. B.; Cicek, N.; Sparling, R.; Berlin, A.; Levin, D. B., Biomass pretreatment: Fundamentals toward application. *Biotechnology Advances* **2011,** *29* (6), 675-685.

39. Jung, Y. H.; H., K. K., Chapter 3. Acidic Pretreatment. In *Pretreatment of Biomass*, Pandey, A.; Negi, S.; Binod, P.; Larroche, C., Eds. Elsevier: Amsterdam, 2015; pp 259-264.

40. Pu, Y.; Hu, F.; Huang, F.; Davison, B. H.; Ragauskas, A. J., Assessing the molecular structure basis for biomass recalcitrance during dilute acid and hydrothermal pretreatments. *Biotechnology for Biofuels* **2013,** *6* (1), 15.

41. Pingali, S. V.; Urban, V. S.; Heller, W. T.; McGaughey, J.; O'Neill, H.; Foston, M.; Myles, D. A.; Ragauskas, A.; Evans, B. R., Breakdown of Cell Wall Nanostructure in Dilute Acid Pretreated Biomass. *Biomacromolecules* **2010,** *11* (9), 2329-2335.
42. Foston, M.; Ragauskas, A. J., Changes in lignocellulosic supramolecular and ultrastructure during dilute acid pretreatment of Populus and switchgrass. *Biomass and Bioenergy* **2010,** *34* (12), 1885-1895.
43. Langan, P.; Petridis, L.; O'Neill, H. M.; Pingali, S. V.; Foston, M.; Nishiyama, Y.; Schulz, R.; Lindner, B.; Hanson, B. L.; Harton, S.; Heller, W. T.; Urban, V.; Evans, B. R.; Gnanakaran, S.; Ragauskas, A. J.; Smith, J. C.; Davison, B. H., Common processes drive the thermochemical pretreatment of lignocellulosic biomass. *Green Chemistry* **2014,** *16* (1), 63-68.
44. Li, C.; Knierim, B.; Manisseri, C.; Arora, R.; Scheller, H. V.; Auer, M.; Vogel, K. P.; Simmons, B. A.; Singh, S., Comparison of dilute acid and ionic liquid pretreatment of switchgrass: Biomass recalcitrance, delignification and enzymatic saccharification. *Bioresource Technology* **2010,** *101* (13), 4900-4906.
45. Selig, M. J.; Viamajala, S.; Decker, S. R.; Tucker, M. P.; Himmel, M. E.; Vinzant, T. B., Deposition of Lignin Droplets Produced During Dilute Acid Pretreatment of Maize Stems Retards Enzymatic Hydrolysis of Cellulose. *Biotechnology Progress* **2007,** *23* (6), 1333-1339.
46. Sun, Q.; Foston, M.; Meng, X.; Sawada, D.; Pingali, S. V.; O'Neill, H. M.; Li, H.; Wyman, C. E.; Langan, P.; Ragauskas, A. J.; Kumar, R., Effect of lignin content on changes occurring in poplar cellulose ultrastructure during dilute acid pretreatment. *Biotechnology for Biofuels* **2014,** *7* (1), 150.
47. Sannigrahi, P.; Ragauskas, A. J.; Miller, S. J., Effects of Two-Stage Dilute Acid Pretreatment on the Structure and Composition of Lignin and Cellulose in Loblolly Pine. *BioEnergy Research* **2008,** *1* (3), 205-214.
48. Pingali, S. V.; Urban, V. S.; Heller, W. T.; McGaughey, J.; O'Neill, H.; Foston, M. B.; Li, H.; Wyman, C. E.; Myles, D. A.; Langan, P.; Ragauskas, A.; Davison, B.; Evans, B. R., Understanding Multiscale Structural Changes During Dilute Acid Pretreatment of Switchgrass and Poplar. *ACS Sustainable Chemistry & Engineering* **2017,** *5* (1), 426-435.
49. Shah, R.; Huang, S.; Pingali, S. V.; Sawada, D.; Pu, Y.; Rodriguez, M.; Ragauskas, A. J.; Kim, S. H.; Evans, B. R.; Davison, B. H.; O'Neill, H., Hemicellulose–Cellulose Composites Reveal Differences in Cellulose Organization after Dilute Acid Pretreatment. *Biomacromolecules* **2019,** *20* (2), 893-903.
50. Muchero, W.; Guo, J.; DiFazio, S. P.; Chen, J.-G.; Ranjan, P.; Slavov, G. T.; Gunter, L. E.; Jawdy, S.; Bryan, A. C.; Sykes, R.; Ziebell, A.; Klápště, J.; Porth, I.; Skyba, O.; Unda, F.; El-Kassaby, Y. A.; Douglas, C. J.; Mansfield, S. D.; Martin, J.; Schackwitz, W.; Evans, L. M.; Czarnecki, O.; Tuskan, G. A., High-resolution genetic mapping of allelic variants associated with cell wall chemistry in Populus. *BMC Genomics* **2015,** *16* (1), 24.
51. Biswal, A. K.; Atmodjo, M. A.-O.; Li, M. A.-O.; Baxter, H. L.; Yoo, C. G.; Pu, Y. A.-O.; Lee, Y. C.; Mazarei, M.; Black, I. M.; Zhang, J. Y.; Ramanna, H.; Bray, A. L.; King, Z. R.; LaFayette, P. R.; Pattathil, S.; Donohoe, B. A.-O.; Mohanty, S. S.; Ryno, D.; Yee, K.; Thompson, O. A.; Rodriguez, M., Jr.; Dumitrache, A.; Natzke, J.; Winkeler, K.; Collins, C.; Yang, X. A.-O.; Tan, L.; Sykes, R. W.; Gjersing, E. L.; Ziebell, A.; Turner, G. B.; Decker, S. A.-O.; Hahn, M. A.-O.; Davison, B. H.; Udvardi, M. K.; Mielenz, J. R.; Davis, M. A.-O.; Nelson, R. S.; Parrott, W. A.; Ragauskas, A. J.; Neal Stewart, C., Jr.; Mohnen, D., Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. **2018,** *36* (3), 249-257.

Chapter 2). Plant cell wall structure and pretreatment

The cellulose, hemicellulose, pectin, and lignin assemble to make the plant cell wall, which serves several critical functions through the life span of the plant.[1] The plant kingdom comprises of unicellular organisms like algae to multicellular flowering plants. Algal biomass is used for advanced 3rd generation biofuel production[2] but the cellulosic 2nd generation ethanol-based biofuel is derived mainly from multicellular plants.[3-4] The multicellular plants grow from seed, and depending upon the number of cotyledons present in the seed, the plants are classified as monocots or dicots.[5] As mentioned in Chapter 2, plants that grow in arid conditions and are not consumed by humans or animals should be selected for 2nd generation biofuel production. There are significant differences in the compositions in the cell wall polymers of monocots which include softwood species like spruce, fir, birch, and dicots which include grasses like switchgrass, sorghum and hardwood species such as poplar.[1] Typical composition of cellulose, hemicellulose, and lignin in softwoods, hardwoods, and grasses are shown in Table 2.1.

These multicellular plants variable from each other in composition and they also have specialized cells with heterogenous composition to carry out unique functions. For instance, the parenchyma cells present in the leaf are responsible for metabolism and food production and have a thin and flexible cell wall. Whereas the sclerenchyma cells are present in the root and have a hard cell wall as they give plant support. Similarly, the xylem vessels also have a rigid cell wall with high amount of hydrophobic lignin as they transport water from roots to leaves.[6]

Table 2.1: Cell wall polymer composition of softwood, hardwood, and grasses

Lignocellulosic material	Cellulose (%)	Hemicellulose (%)	Lignin (%)
Softwoods	45–50	25–35	25–35
Hardwoods	40-55	24-40	18-25
Grasses	25–40	35-50	10-30

One of the reasons for variations in the compositions of the cell wall polymers is for controlling the rigidity of the cell wall and helping the cell perform various other functions like turgor pressure,[7] giving shape and strength to the plant cell, building interactions with the microbes, and protection against potential pathogens.[8-9] The typical chemical and molecular characteristics of cell wall polymers are described below.

2.1). Individual polymers of the plant cell wall

2.1.1). Cellulose

Cellulose is a linear polymer that consists of D-glucopyranosyl residues connected by β (1→4) glycosidic linkages.[10-11] This homopolysaccharide has cellobiose as the repeating unit, and the structure of one glucan chain is shown in Figure 2.1.[12]

2.1.1.1). Cellulose fiber formation and general aspects of cellulose structure

Plants, some bacteria such as *Acetobacter xylinus* and algae, synthesize and secrete cellulose.[14] These organisms have an enzyme complex, cellulose synthase embedded in the plasma membrane, that synthesizes and extrudes glucan chain.[15-16] Both inter and intra-chain hydrogen(H) bonding occurs due to the glucan chain being linear and the hydroxides being evenly distributed on both sides of the glucopyranosyl residues.[13] While mentioning H-bonding, it is important to discuss the configuration of C6 hydroxymethyl group.[12] Early modeling studies show that the hydroxylmethyl group on the C6 of glucopyranosyl residue in cellulose I is in the trans-gauche (tg) configuration. This is based on the comparison of models of cellulose I with the three configurations (gg = gauche-gauche, tg = trans-guache, gt = gauche-trans) to x-ray diffraction data.[13]

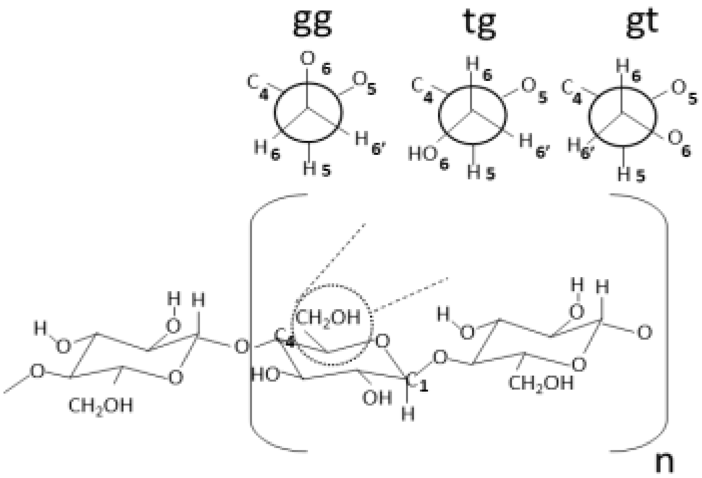


Figure 2.1: Single-chain of glucan of cellulose in which the glucopyranose unit is in thermodynamically stable chair form and connected by β (1→4) linkage. The three most probable positions of hydroxymethylgroup (gg = gauche-gauche, tg = trans-guache, gt = gauche-trans) around the C(6) are also shown[12-13]

The neutron diffraction data and x-ray synchrotron fiber diffraction data show that cellulose I (from plant or algal source) does have all hydroxymethyl groups adopting the tg configuration.[17-18] However, a recent study shows that the surface glucan have gg configuration, the inner core chains have gt configuration, while the chains in between contributing to the crystalline cellulose I are in tg configuration.[19] The C6 hydroxylmethyl group participates in forming hydrogen bonds both within the glucan due to which individual chains have linear structure and between chains that introduce order or disorder in the cellulose fiber. The order or disorder is dependent on the regularity of occurrence of the intrachain H-bond. Eventually both the inter and intra-chain H-bonds and the Vander Waal forces are responsible for the structure of cellulose I (Figure 2.2).

The glucan chain coalesces to form cellulose microfibrils.[14-16] However the number of glucan chains present in a single microfibril in plant cellulose is debated, but through several different studies it is proposed to be either 18[23-24], 24[25], or 36.[21-22] Microfibrils are of long in length as compared to its cross-sectional size (approximately 3-5nm); however, the exact area of the microfibrillar cross-section and its shape are not yet determined. Several models of the cross-sectional shape exist as shown in Figure 2.3. The shape is important to know because that would determine the hydrophobic and hydrophilic surfaces of the cellulose microfibril.[25-26]

The microfibrils coalesce together to form the macrofibril or the cellulose fiber .[14-16] The glucan chain that are extruded from bacteria and algae usually have around 50 to 160 chains being secreted from one cellulose synthase complex which coalesce to form highly crystalline cellulose.[14]

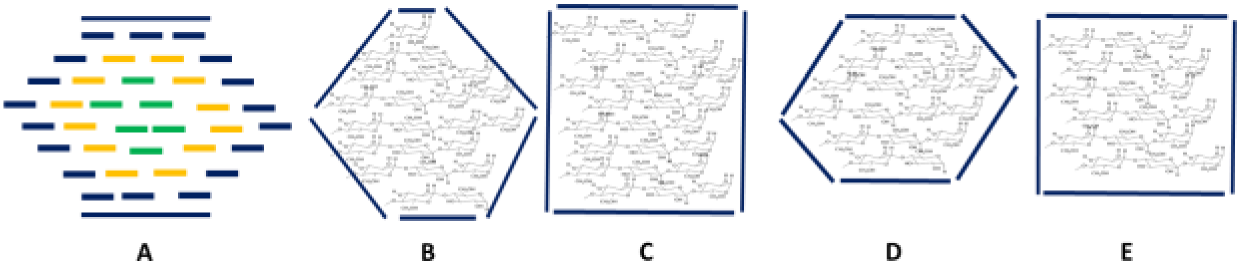


Figure 2.2: Cross-section of cellulose microfibril with 36 glucan chain in which the core (green) chains are not as flexible as the surface glucan chains (blue). While this is the most common depiction found in literature,[20,] there is growing evidence that the microfibril has fewer chains than 36[21-22]. Along the length of the microfibrils, the glucan chains are present in planar arrangement and have intra and inter hydrogen bonds.

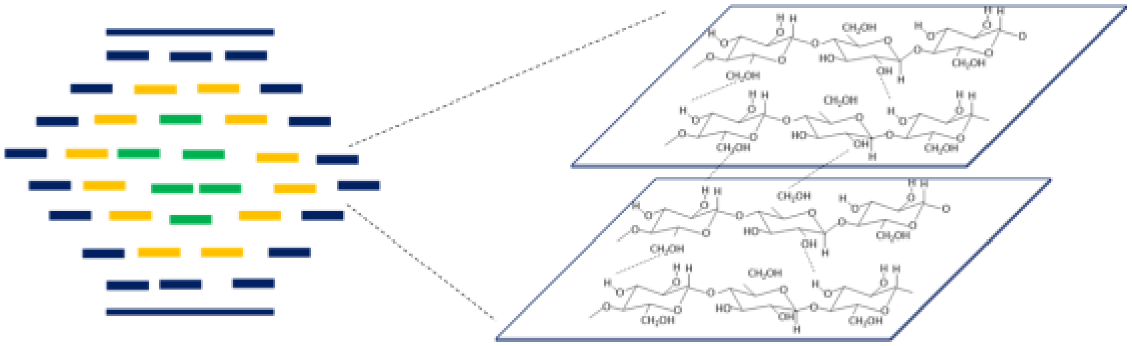


Figure 2.3: The probable shape of the microfibril cross-section based on the number of glucan chains. Model A has a hexagonal shape with 36 glucan chains, Model B and C both have 24 chains, but the hydrophobic surface in model B is much shorter than that in model C. Model D, and E consist of 18 glucan chains each.[21]

Besides differences in the number of glucan chains, the atomic-scale structure details showed differences in the arrangement of the glucan chain with respect to one another between bacterial and algal and plant cellulose. NMR studies show that the packing of the glucose molecule in the microfibril occurred in two forms: Iα and Iβ.[27-28] In both these forms the glucan chains are parallel and the reducing ends were aligned in the same direction.[17-18] The difference between them was not known until the X-ray crystal structure was solved.[17] For cellulose Iα the unit cell is triclinic while containing one glucose monomer per unit cell, while for cellulose Iβ it is a monoclinic unit cell containing two glucose monomer per unit cell. Due to this for cellulose Iβ, the glucose molecules in chains parallel to each other stack on one another while for Iα, the adjacent glucan chains were at an offset of ½ c axis. Cellulose Iα is meta-stable, and it can be converted to the thermodynamically more stable cellulose Iβ by annealing.[29] The differences between their glucan chain arrangement are highlighted in Figure 2.4.[17-18]

2.1.1.2). Cellulose crystallinity

Cellulose is semi-crystalline as it has both crystalline and the non-crystalline or amorphous region.[30-31] Depending on the organism from which the cellulose is extracted, the amount of crystalline to amorphous region varies. The term crystallinity index was coined, and it is the measure of the relative amount of crystalline material in cellulose.[31] Different techniques such as infrared spectroscopy,[32] x-ray diffraction,[33-34] nuclear magnetic resonance (NMR)[35-36] are used to measure the crystallinity index. Crystallinity determination from X-ray diffraction depends on the way the amorphous phase is determined (detailed description for calculation shown in Chapter 4).

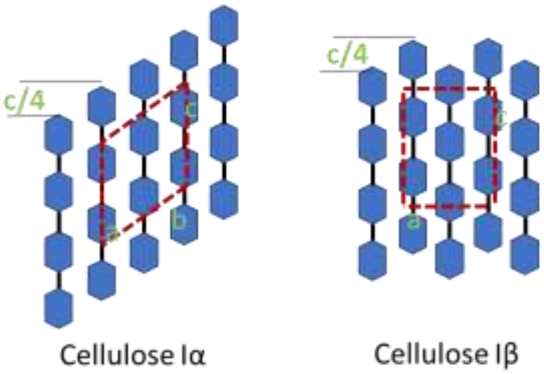


Figure 2.4: The glucan chains are at ½ axis offset between glucan chain and has a monoclinic unit cell in cellulose Iα whereas the cellulose Iβ has glucan chain parallel have a triclinic unit cell[12, 17-18]

Each technique gives variable results and comparisons of crystallinity index between different cellulose samples if made, must have the crystallinity index measured by that particular technique.[31] Table 2.2 shows a comparison of crystallinity index estimated using different techniques for commercially available cellulose.

2.1.1.3). Degree of polymerization of the cellulose

The glucan chain from different cellulose sources show a variation in their length.[39] The length of the cellulose fiber is quantified by degree of polymerization (DP). It can be measured as number-average DP (DPn), weight-average DP (DPw), or viscosity-average DP (DPv).[40] Gel permeation chromatography (GPC) is the most commonly used technique of measuring cellulose DP. Both DPn, DPw can be determined with GPC, and also the polydispersity index (PDI=Mw/Mn) can be calculated. This value can be used to determine the distribution of the polymer molecular weight.[41] In a typical procedure,cellulose is first derivatized using tricarbanilation reaction in which reaction of cellulose with phenyl isocyanate in pyridine takes place. After this, the reaction is quenched with methanol, and cellulose tricarbalinate is precipitated.[42] This can then be dissolved in tetrahydrofuron and the DP measurements be made with GPC column. The column is calibrated with standards with known and varying molecular weights such as polystyrene.[41]

The cellulose viscosity-average DP (DPv) can be determined by viscometry, and the values are relatively quickly and conveniently obtained as compared to GPC. But they are not reliable as no clear information concerning the molar mass distribution can be determined. Additionally, metal solutions are used in viscometry which can degrade cellulose.[40, 42] Despite all the drawbacks, viscometry techniques are more commonly used for DP determination of cellulose from lignocellulosic biomass due its complex nature. [32]

Table 2.2: Percent crystallinity calculated with different techniques and method of hardwood, softwood, and grasses

Feedstock	XRD[37]			Feedstock	NMR[38]
	Peak height	Amorphous Subtraction	Peak deconvolution		C4 peak separation
Corn Stover	47	39	37	Switchgrass	44
Norway Spruce	47	56	33	Pine	63
Hemp fiber	77	49	60	Poplar	63

The DP of cellulose (measured using viscometric method) from various hardwoods (such as poplar) are found in range of around 1,500 - 4,500, 3,500 - 4,500 for softwood, and 1,800 to 4,000 for agricultural residues.[40]

2.1.1.4). Treatment with chemicals and change in cellulose crystalline allomorph

The atomic structural location of the glucose in the cellulose fiber can be changed by treatment with chemicals such as sodium hydroxide, ammonia or ethylene diamine and glycerol.[12] The treatment done to cellulose Iα [18]and Iβ [17]results into cellulose II, III, or IV.[12, 27, 30, 43] Further details of the treatment are shown in schematic Figure 2.5.

Cellulose II is made by dissolving cellulose I into concentrated alkali (>10% sodium hydroxide) and regenerating by precipitating the dissolved mixture with water.[44] In cellulose II the hydroxymethyl carbon is in the gt position unlike cellulose I where it is in the tg position.[12] This form of cellulose II is the most thermodynamically stable state of cellulose. The chains are arranged in antiparallel manner. A bifurcated hydrogen bond is formed at O3-H···O5 as the major component and O3-H···O6 as the minor component.[27, 45-46] Cellulose III has a confirmation similar to cellulose II however, the chains are parallel. Cellulose III is made by soaking cellulose I with liquid ammonia or with organic amine such as ethylenediamine. It also has the hydroxymethyl group in the gt position and bifurcated hydrogen bond at O3-H···O5 and O3-H···O6 forms.[47] The existence of cellulose IV is debated in literature however some studies show it has the highest number of H-bonding indicating a very stable structural form of cellulose.[48-50]

Single crystals of cellulose IV_{II} grow at higher temp than cellulose II which is the most stable form of cellulose.[49] The crystallite length is shorter than cellulose I. However, through moisture regain experiment it was found that it is more accessible than cellulose I.

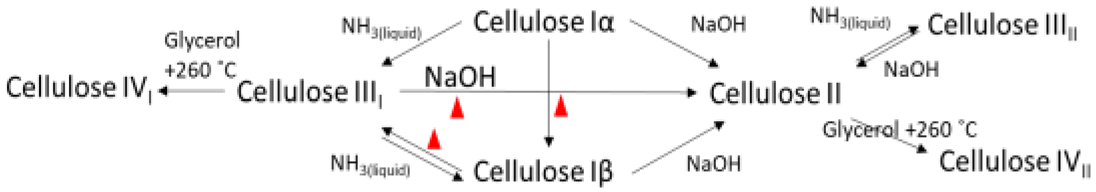


Figure 2.5: Conversion of cellulose I to Cellulose II, III, and IV by treating cellulose I with different chemicals. The red triangle is to denote heat needed for the conversion. The single arrow indicates irreversible reaction while the double arrow indicates a reversible reaction[12]

The tensile property and strength is comparable to cellulose I even though DP is shorter than cellulose I.[50]

2.1.2). Lignin

Lignin is a rigid three-dimensional network mostly made of phenyl propane structures (C3C9) and is found in cell walls of higher plants and grasses.[51] The hydrophobic lignin present in the vascular tissue of plants helps carry water from roots to different parts of the plant. Lignin biosynthesis starts by the conversion of the amino acid, L-phenylalanine or L-tyrosine to coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol by series of enzymatically driven steps. Figure 2.6 shows the main intermediates and pathways involved in this conversion.[51-53]

These monolignols become part of the lignin as phenylpropanoids groups and are denoted as courmaryl (H), gauicyl (G), and sinapyl (S).[52] Figure 2.7 depicts the structure of the phenylpropanoids groups. Complete elucidation of lignin structure is challenging because lignin cannot be extracted by solvents which do not change the lignin structure.[54] Moreover, the cell wall native lignin is associated with polysaccharide like hemicellulose, pectin and small molecules like tannin, which contribute to structure determination complexity in situ.[55-56] Various chemical degradation reactions of lignin such as nitrobenzene oxidation, ozonolysis, thioacidolysis have been used to determine monolignol composition/content and linkages. These methods are laborious so they are replaced by structural characterization techniques.[57-58] Nonetheless, combination of chemical isolation[59] techniques and structural characterization using spectroscopy, Nuclear Magnetic Resonance (NMR),[60-63] UV-microscopy,[64] Coherent Anti-stroke Raman Scattering (CARS),[65] Infrared (IR) and Raman Spectroscopy,[66] Electron Paramagnetic Resonance (EPR) offer information about isolated lignin structure.[67]

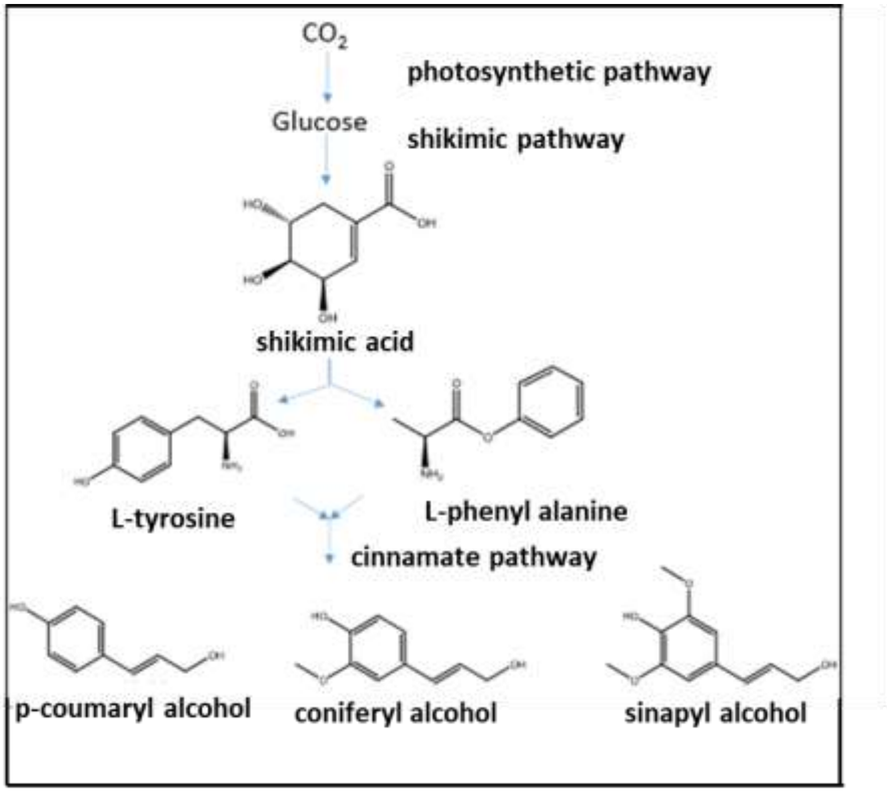
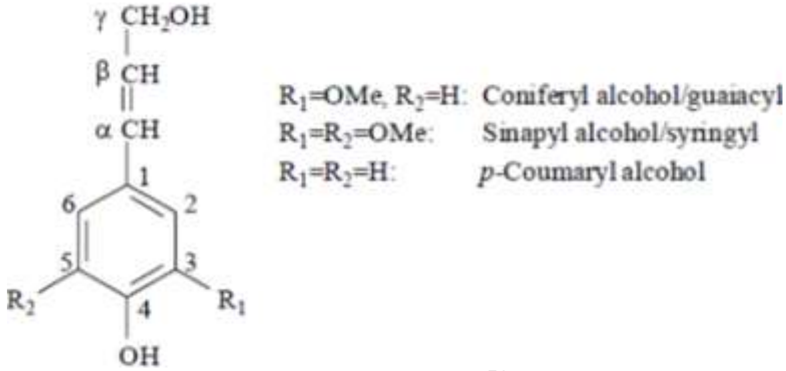


Figure 2.6: Schematic showing the main intermediates and pathways involved in the formation of phenylpropanoids in plants[51]



R_1=OMe, R_2=H: Coniferyl alcohol/guaiacyl
R_1=R_2=OMe: Sinapyl alcohol/syringyl
R_1=R_2=H: *p*-Coumaryl alcohol

Figure 2.7: Nomenclature of the monolignol skeleton[52]

Based on structural studies and mutant plant studies,[60, 68] predictions about the native lignin structure can be made. Milling of wood imparts changes in the native lignin structure however changes are less than any other technique used to isolate lignin.[64] It has been assumed that 30-50% lignin obtained from milled wood lignin represents a structural average of the total lignin.[69] Due to variable interlinkage formation, lignin is highly branched polymer with weight average of molecular weight (Mw) ~20,000 in softwood and polydispersity of 2.5 while the molecular weight is slightly in hardwood, along with a polydispersity of 3.5.[70]

The ratios of monolignol in plant species are variable (Table 2.3), and this leads to the formation of a heterogeneous and rigid lignin structure.[51-52] Lignin is very important for the plant as it imparts strength and protects against pests.[51-52] The coupling of monolignols is via the formation of free radicals. On a single monolignol there are multiple sites for a radical form and be available for bond formation, as shown in Figure 2.8. Due to this, many interunit linkages such as resinol (β-β), phenylcoumaran (β-5), biphenyl (5-5), and 1,2-diaryl propane (β-1) are possible.[52-53, 55] Due to differences in the monolignol composition, the occurrence of the linkage type also differs based on the type of wood (Table 2.4).

2.1.3). Hemicellulose

Hemicellulose comprises of a board class of polysaccharides which are non-cellulosic and extracted from the cell wall. The backbone of the hemicelluloses contains monomeric sugars such as glucose, mannose or xylose joined β (1-4) linkages. Most hemicellulose comprises of side chains that vary in length, type of sugar monomers, sometimes have uronic acid residue and degree of acetylation. Based upon the sugar backbone the hemicelluloses are divided into four categories namely xyloglucan, xylans, mixed linked

Table 2.3: Composition of lignin and the G/S/H lignin ratio that make the lignin[52-53, 71-73]

Feedstock	Lignin Content (%)	Guaicyl (H) (%)	Sinapyl (S) (%)	p-coumaryl alcohol (H) (%)
Grasses	12-18	40	50	10
Softwood	27-30	90	10	-
Hardwood	18-26	50	50	-

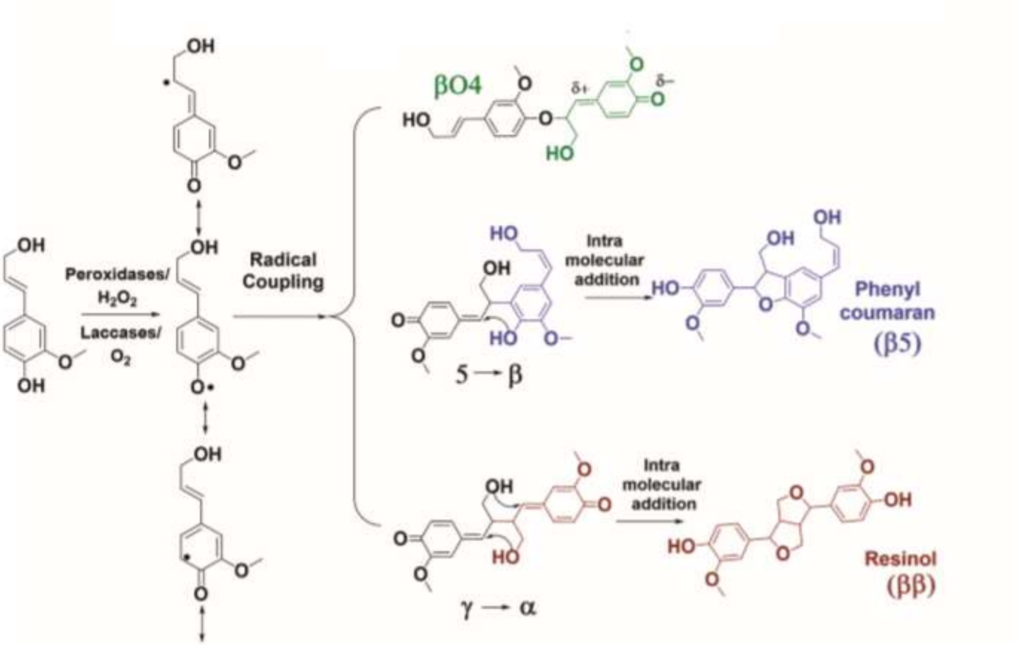


Figure 2.8: Enzymes such as peroxidase or laccases in the presence of hydrogen peroxide convert the monolignol to a free-radical which then couples with another monolignol forming various interunit linkages to for dignols[55]

Table 2.4: Type of linkages between monolignol in softwood and hardwood ((number of linkages per 100 C9 units)[52, 61, 71, 74]

Type of wood	β-O-4	β-5	β-β	5-5	β-1	Dibenzodioxocin
Softwood	45-50	9-12	2-4	19-22	7-9	5-7
hardwood	60-62	3-11	3-12	3-9	1-7	0-2

glucans and mannans.[75-76] The typical amounts found in primary and secondary cell wall are shown in Table 2.5

2.1.3.1). Xyloglucans

Xyloglucans are found in the primary cell wall of all higher plants but in variable amounts, as shown in Table 2.4.[75] They are often directly released from primary cell wall by endoglucanase treatment. However, extraction with alkali followed by endoglucanase treatment to form xyloglucan oligosaccharides have also been reported.[81] The molecular weight of xyloglucans can be as high as 200 kDa.[27] Determination of structural features such as repeating backbone monomer patterns, linkages, and side-chain monomers of xyloglucan polysaccharide is done by studying xyloglucan oligosaccharides.[82-83] The xyloglucan from hardwood and softwood have a β(1-4) linked glucan backbone with many side chains. Xylose containing side chains are connected to the glucose in the backbone by α(1-6) linkages.[82-83] The known and commonly occurring fragments are shown in Figure 2.9.

2.1.3.2). Xylans

Xylans, in general, are a type of hemicellulose that contains a β-1,4-xylosyl backbone. Xylans are grouped into four major types: O-acetylglucuronoxylan (AcGX), arabinoglucuronoxylan (AGX), O-acetylglucuronoarabinoxylan (AcGAX), and O-acetylarabinoxylan (AcAX).[78] However, it should be noted that the structural characteristics like the molecular weight, the type of side chains, degree of acetylation is different between plant species and cell types, and dependent on the growth of the plant.[78, 86] The presence of commonly observed groups, linkages and other structural features in bioenergy related feedstock are shown in Figure 2.10.[87]

Table 2.5: Composition of the different hemicellulose in the primary and secondary cell wall of hardwood, softwood, and grasses[75-76]

Polysaccharide	Hardwood		Softwood		Grasses[77]	
	Primary	Secondary	Primary	Secondary	Primary	Secondary
Xyloglucan	20-25	Minor	2-5	Minor	10	-
Glucoronoxylan[78]	-	20-30				
Glucoronoarabinoxylan[78]	5		20-40	40-50	2	5-15
(Gluco)mannan[79]	3-5	2-5	2	0-5	-	-
Galactoglucomannan[80]	-	0-3				10-30
Mixed link glucan	-	-	2-15	Minor	-	-

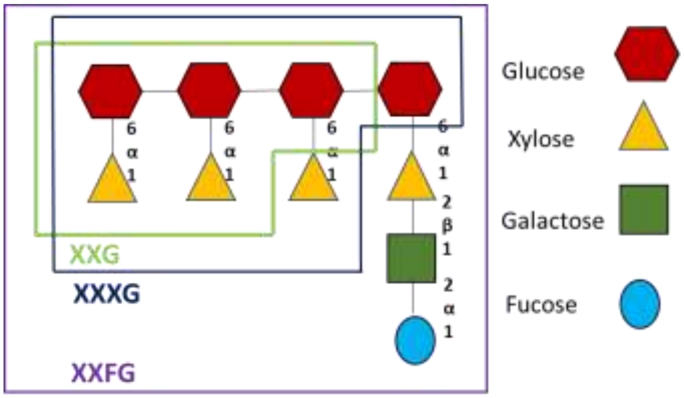


Figure 2.9: Common xyloglucan oligosaccharide fragments found mostly in dicots are nonasaccharide (XXFG), heptasaccharide (XXXG)[84] and pentasaccharides (XXG)[81, 83, 85]

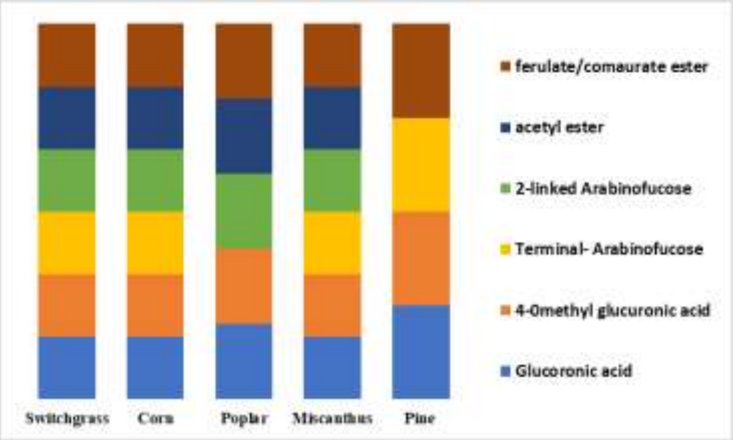


Figure 2.10: Salient structural features of the different xylans present in grasses (switchgrass, miscanthus, and corn), hardwood (poplar) and softwood (pine)[87]

The AcGX are mostly detected in the secondary walls of the conducting vessels like xylem and phloem. They are monodisperse in length, and every tenth xylosyl residue is substituted at O-2 position with (4-O-methyl)-α-d-glucuronic acid((Me)GlcpA).[88] Additionally, these polysaccharides are heavily O-acetylated at either O-2 or O-3 positions of the xylosyl residue making about 34-49% of xylosyl residues acetylated.[89]

Arabinoglucoronoxylan (AGX) are present in softwoods although in minor amounts. They are substituted, on average, with two 4-O-methylα-d-glucuronic acid groups at O-2 and one α-l-arabinofuranose residue at O-3 per every ten xylose units. They are highly decorated with side-chain substitutions but are not O-acetylated except for the AGXs from some species of softwoods.[90]

Arabinoxylan is present in the primary cell wall of grasses. About 30% of the primary cell wall of grasses have arabinoxylan whereas the primary cell wall of dicots and monocots have less than 5%.[91] Both AcGAXs and AcAX are acetylated like the AcGX, but the extent of acetylation is low.[87]

The acetyl groups are attached to the backbone xylosyl residues. There are also the arabinofuranose substituents, which could also be acetylated at O-2 position. Most importantly the O-5 position of the arabinofuranose residues is involved in forming ester bonds with the ferulic or p-coumaric acids. Through such bonds it is shown that lignin could bond to xylans in grasses.[56, 77]

2.1.3.3). Mannans

Mannans appear to have been very abundant in early land plants and are still abundant in mosses and lycophytes.[79-80] Besides the presence of xylans, the secondary cell wall of softwoods also contains mannans in the form of galactoglucomannans or glucomannan.

These mains have a β-(1→4)-linked glycosyl residues backbone containing both glucose and mannose. The ratio of glucose to mannose can vary between 1:3 or 1:4. Mannans and galactomannans have a complete mannose backbone.[80, 92] The mannans and glucomannans are often acetylated. Mannans have been studied for their role as seed storage compounds. They play essential roles, as mannan synthase knockdown arabidopsis mutant did not grow.[79, 93]

2.1.4). Pectin

2.1.4.1). Homogalacturonans

Homogalactuonans (HG) are formed from galacturonan connected through α (1-4) linkage.[94] It is methyl esterified, however the degree that it can vary based on the plant species, cell type, and age of the plant. Nonetheless, the methylesterification of homogalacturonan is shown to be as high as 80%.[95] The non-esterified part of HG is involved in gel formation within the cell wall. The gels are formed when calcium ions crosslink adjacent non-esterified HG regions of two chains (Figure 2.11). The non-esterified HGs are mostly found in the cell corner and middle lamella region. Highly methyl esterified HGs also take part in the gel formation.[96] The gels, in this case, are shown to form in a water-deficient environment, and the chains are held together by hydrogen bonds and hydrophobic interactions. The esterified pectins are present throughout the primary cell wall but mostly surround the cellulose hemicellulose networks.[21, 95, 97]

2.1.4.2). Rhamnogalacturonan I

RG-I has a backbone made of altering repeat units of rhamnosyl and galactosyluronic acid residues.

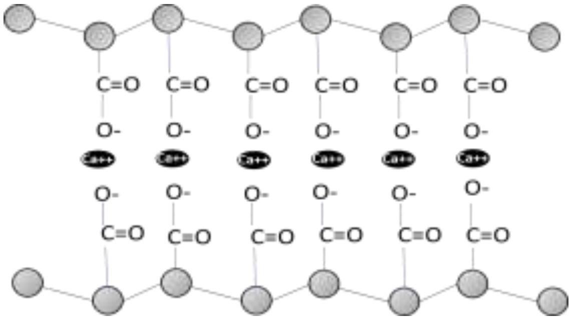


Figure 2.11: Calcium bridges are proposed to cross-link adjacent homogalacturonan chains[98-99]

The backbone repeating unit is connected as (→4-α-D-GalpA-(1→2)-α-L-Rhap-(1→) It has side chains that differ from each other and are connected to the backbone through the C-4 of rhamnosyl residue.[100]The side chains contain arabinosyl, galactosyl, and minor amounts of fucosyl and glucosyluronic acid residues.[101] A single side chain can have about 30 sugar residues.[102] RG-I content results in about 5-10% total cell wall content in hardwood, 4-7% in softwood and about 1% in grasses. While RG-I is easily extracted after treating the cell wall with hydrolytic enzyme-like endo-polygalacturonase, sodium carbonate solubilization prior to enzyme hydrolysis is also at times done to extract RG-I. [103] The molecular weight of Rg-I ranges between 10^5 to 10^6 Daltons, and as it elutes as a broad peak after size exclusion chromatography, it can be concluded that RG-I is polydisperse.[97, 100]

2.1.4.3). Rhamnogalacturonan II

Out of all known plant polysaccharides, the structure of RG-II is the most complicated.[104] The backbone contains the D- galactosyluronic acid residues connected by α(1→ 4) linkages. Some of the residues are esterified with methyl groups. The backbone found to be only 30 residues long; however, the polymer is highly branched. While most polysaccharides are singly branched, RG-II has at least 6% of the total backbone residues doubly branched.[105] Additionally, RG-II forms dimers by co-ordination with boron. Other heavy metals like lead and strontium have also been detected with RG-II.[106-107] Unusual glycosyl residues such as 3-deoxy-D-manno-2-octulosonic acid and 3-deoxy-D-lyxo-2-heptulosaic acid have also been detected in RG-II. Apiosyl, 2-O-methyl fucosyl, 2-O-methyl xylosyl, arabinosyl are few the other key monomers found in RG-II. Apiosyl is shown to participate in formation of borate bridges.[106-107] RG-II is complex and plant has to spend high amount of energy to make this polysaccharide. Due to this, it is usually found in low amount. The complexity of the structure makes it difficult to be broken down by microbial enzymes and hence provides the plant protection against microbes.[98]

2.2). Structure of plant cell wall

While the compositional heterogeneity imparts plant cell walls with many diverse functions, studying the structure of the cell wall becomes very challenging. There are many unanswered questions about the structure of the cell wall.[21] These include conformation of the cell wall polymers in the plant, interaction between them, their biosynthesis and assembly into the cell wall.[108-111] These are very important to study as a better understanding would help reduce the recalcitrance to enzymatic hydrolysis. Besides producing biofuels efficiently, the structure of cell wall is also of interest for separating cell wall polymers for conversion to bio-products,[67] extraction of biologically active small molecules and enzymes and extraction of polysaccharides like pectin, hemicellulose[75-76] that are used in food industries as stabilizers, gum, and gels.[94] The deposition of the cell wall polymers occurs layer by layer during the growth and development of the cell wall.[109] This makes the cell wall polymer composition to be dynamically changing as the plant cell matures.[112] The deposition of polysaccharides and lignin in the cell wall from cell formation until cell death is shown in Figure 2.12. Initially, during the growth phase of the cell, the wall is a part of apoplast. Apoplast is the space between the plasma membrane and the neighboring plant cell. The wall that surrounds the growing cell and is capable of growth and expansion is called the primary cell.[7] Later, as the plant cell continues to develop, different hemicelluloses start filling the space in the cell wall.

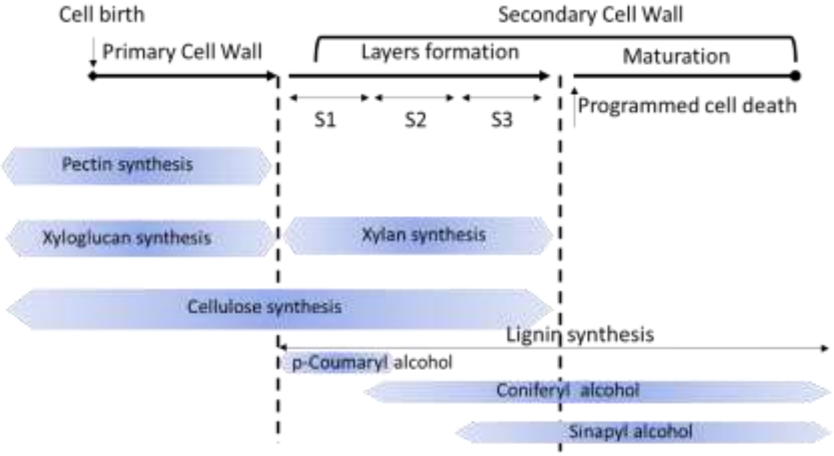


Figure 2.12: The schematic depicts the formation of different cell wall layers as the new cell forms and matures. During the formation of each layer, different polymers get synthesized and deposited into the cell wall.[112-114]

There about two thousand genes that get transcribed into glycosyl transferase and glycosides to make the different hemicelluloses.[115] Due to this, there is a huge diversity and complexity in the structure of the hemicelluloses and pectins of the cell wall. The region between two adjacent cells, where the primary cell wall of each cell meet is called the middle lamella. The secondary cell wall is found in specialized cells like xylem vessels and fiber cells. The secondary cell wall usually forms after the cell ceases to grow.[7, 116] Figure 2.13 shows the primary, secondary and middle lamella region of the cell wall. All differentiated plant cells contain the primary and secondary cell wall, but the volume occupied by each of these cell wall layers is variable.

2.2.1). Primary cell wall

The major component present in primary cell walls isolated from higher plant tissues and cells have cellulose (14-50%) and other polysaccharides like pectin (30-50%) and xyloglucan (20-30%). Structural glycoproteins (hydroxyproline-rich extensins), phenolic esters (ferulic and coumaric acids), ionically and covalently bound minerals (e.g. calcium and boron), and proteins are an integral part that binds the polymers together but are found in lesser amounts.[118] The cell wall is formed during the growth phase of the cell, and a class of proteins called expansins are responsible for regulating the expansion of the cell wall.[117] The texture of fruits and vegetables is modulated by primary cell wall. [117]

The organization of the primary cell wall polymers and other components that make the three-dimensional cell wall architecture is a current major area of research. Different models such as the tether model,[7, 119] the diffuse layer model[120] and the stratified layer model[121] have been proposed that explain the cellulose, pectin, and xyloglucan arrangement in the primary cell wall with respect to each other (Figure 2.14).

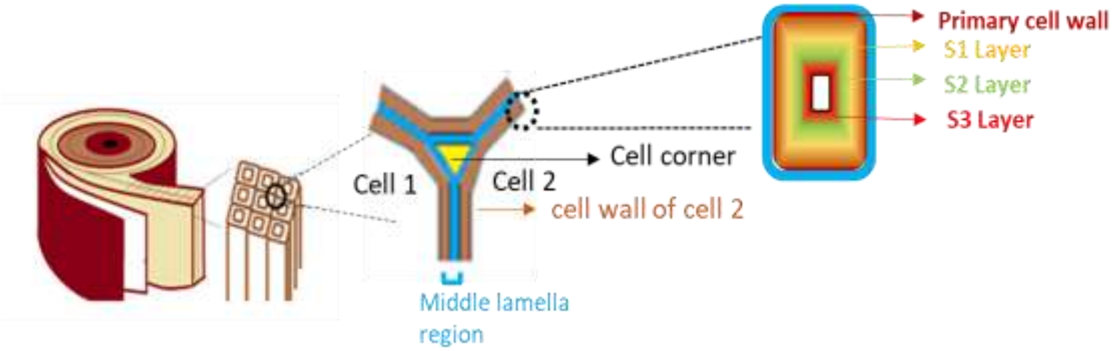


Figure 2.13: Schematic showing the cell wall adjacent to two cells. The cell wall of each cell is made of primary and secondary cell walls. The secondary cell wall has S1,S2,S3 layers.[117]

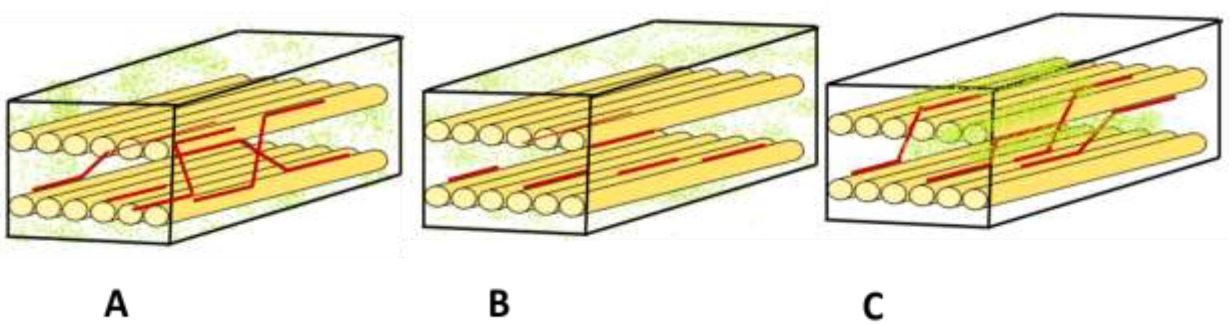


Figure 2.14: Proposed primary cell wall models A). Tethered model where xyloglucan (red) cross-links the cellulose microfibrils (yellow), and pectin (green) form a network, B). Diffuse model wherein the xyloglucan coat the cellulose surface, and they are present in pectin network. C). Stratified layer model in which xyloglucan crosslinks the cellulose microfibrils, but the pectin is present as strata separating the cellulose-xyloglucan layers[119-122].

According to the tethered model, two networks form a mesh-like structure. One of the networks is formed from pectin while the other comprises of xyloglucan and cellulose that are bonded together by non-covalent interactions such as hydrogen bonds and cross-link the cellulose microfibrils.[122] The other model is the stratified layer model in which the xyloglucan is also hydrogen-bonded and cross-links cellulose microfibrils. However, the arrangement of xyloglucan and cellulose differs from the tethered model.[121] The xyloglucan-cellulose are depicted to exist as lamella, and the lamella are shown to be separated by the pectin network. Finally, in the diffuse model, there is no crosslink between xyloglucan and cellulose, but the cellulose and xyloglucan are connected only through hydrogen bonds.[119] The xyloglucan and cellulose are embedded in a network formed from pectin. The evidence to support xyloglucan-cellulose bonding comes from extractability studies of xyloglucan from plants tissue[81, 123] and in vitro binding studies.[124] Additionally, several microscopy studies show xyloglucan to be present in the spaces between microfibrils.[119] The role of the tethered xyloglucan has been proposed to limit cell enlargement by holding the cellulose microfibrils together in the lateral direction.[122] However, the tethering of cellulose by xyloglucan in the model does not explain some of the experimentally observed results. For instance, no difference in deformation of the plant tissue in which xyloglucan was broken down by xyloglucan specific endoglucanase was observed when creep force was applied.[125] Additionally, no evidence supporting cellulose xyloglucan interactions leading to network formation was found with solid-state NMR.[126] Rather pectin and cellulose interacting to form a network-like structure have been reported with solid-state NMR.[127] Based on these finding, cross-links between xyloglucan and cellulose may not exist. While the role of xyloglucan for expansion of cell wall has been negated, xyloglucan is thought to provide mechanical strength by acting as junctions that connect the cellulose fibers. Such connecting points are termed as bio-mechanical hot spots of the primary cell wall.[21] Cell walls have cellulose microfibrils that deposit to form a lamella. The primary cell wall has lesser number of lamellae as compared to the secondary cell wall.[119] The pectin network is shown to limit cellulose accessibility.[128]

2.2.1.1). Polymer interactions

Cellulose pectin interactions

In vitro studies have reported pectin binding to cellulose by hydrogen bonding.[124] This binding is much weaker than that observed for xyloglucan and cellulose. Bacterial cellulose synthesized in the presence of pectin shows that pectin does bind to cellulose.[129] Extracts from various plants showed the pectic polysaccharide remains attached to the cellulose in unextracted residue.[55] Most models of the primary cell wall show separate cellulose-xyloglucan network and pectin network and interactions between these networks due to entanglement. However, with progress in NMR instrument development, cross-peaks indicative of pectin side chains interacting with cellulose has been observed. The cross-peaks remain in a partially homogalacturonan cell wall suggestive of pectin being present between cellulose microfibrils.[127]

Different pectic interactions in the pectin network:

The pectic polysaccharides identified in the primary cell wall include homogalacturonan (with different degrees of methyl esterification), rhamnogalacturonan I and rhamnogalacturonan II.[104] Pectin polymers are extracted from a variety of plants by using mild chemicals, hot water, chelators, and even cold water.[130] However, isolating them with different chromatographic separation methods has not been possible unless the extracted pectin containing sample has been treated with glycosidase. This shows that the different pectins are connected to each other covalently by glycosidic bonds. Additionally, different degree of methyl esterification of pectins and chelation of pectin with boron and calcium resulting in network formation is also known.[106-107] Homogalacturonan is linear but they align and cross-link to form network. The mechanism by which they form networks is not completely known, but methyl esterified homogalacturonan can interact with other methyl esterified pectins by forming hydrogen bonds or hydrophobic interactions.[96] A divalent cation such as calcium is shown to participate in pectin network formation by forming ionic interactions with the carboxylic group of the uronic acid present in galacturonic or glucuronic acid of the pectin polysaccharide.[94] Borate bridges formed by interaction between boron and ester linkage with RG-II are also shown to form. Further, it is shown that boron couples with two rhamnogalacturonan II forming dimers.[107]

2.2.2). Secondary cell wall

Eighty percent of the cell wall of grasses, hardwood and softwood feedstock considered for biofuel production is the secondary cell wall.[131] These types of cell walls have been a critical step in the evolution of land plants. The secondary cell wall is present in conducting tissue cells that are responsible for transportation of water and nutrients from ground to different parts of the plant. Conduction of water and nutrients is made possible due to the presence of lignin in the secondary cell wall of these tissues. The lignified cell wall provides mechanical strength to plant due to which the plant stands upright and also provides the cell wall rigidity.[116] The secondary cell wall is further divided into S1, S2 and S3 layers (Figure2.13), which have different extent of lignification in them. The secondary cell wall has different hemicelluloses type than found in the primary cell wall. Depending on the type of woody materials, hemicelluloses such as arabinoxylan, xylans, glucuronoxylan, and glucomannans are present in the secondary cell wall.[75] The orientation of the lamella that contains the cellulose microfibrils in S2 layer is known traverse to S1 and S3 layer. This type of orientation of the lamella gives additional strength to the cell wall.[132] While Table 2.1 at the start of this chapter shows overall cell wall polymers composition, Table 2.6 shows a comparison of polymer composition between primary and secondary cell walls of softwood and hardwood. Cellulose, hemicellulose and lignin organization is shown in Figure 2.15.

2.2.2.3). Polymer interactions

Cellulose hemicellulose interactions

Both hemicellulose and cellulose have a linear backbone in which adjacent sugar monomers are connected by equatorial glycosidic linkage at C1 and C4 positions. Due to this, non-covalent interactions such as hydrogen bonds are possible between hemicellulose, and cellulose.[75] Evidence in support of such interactions came when cell walls were treated with mild to harsh chemicals sequentially. After treatment with harsh chemicals such as 4 M potassium hydroxide, the residual wall showed the presence of sugars from cellulose and hemicellulose.[130] Additionally, treatment of the plant cell wall with strong chaotrope such as 4-methyl-morpoline-N-oxide hydrate can solubilize both hemicellulose and cellulose.[133] The binding affinities of different hemicellulose on cellulose surface have been measured in vitro, and it has been shown that xyloglucan strongly binds to cellulose.[124] Further, changes in the cellulose structural organization has been observed when bacterial cellulose was synthesized in presence of various kind of hemicellulose.[134]

Table 2.6: Distribution of polymers in the primary and secondary cell wall of hardwood and softwood[103]

Polymer	Cell wall (wt%)		
	Primary	Secondary	
		Hardwood	Softwood
Cellulose	20-30	37-57	38-52
Lignin	0	17-30	26-36
Pectin	30-35	<10	<10
Hemicellulose	25-30	20-57	16-27

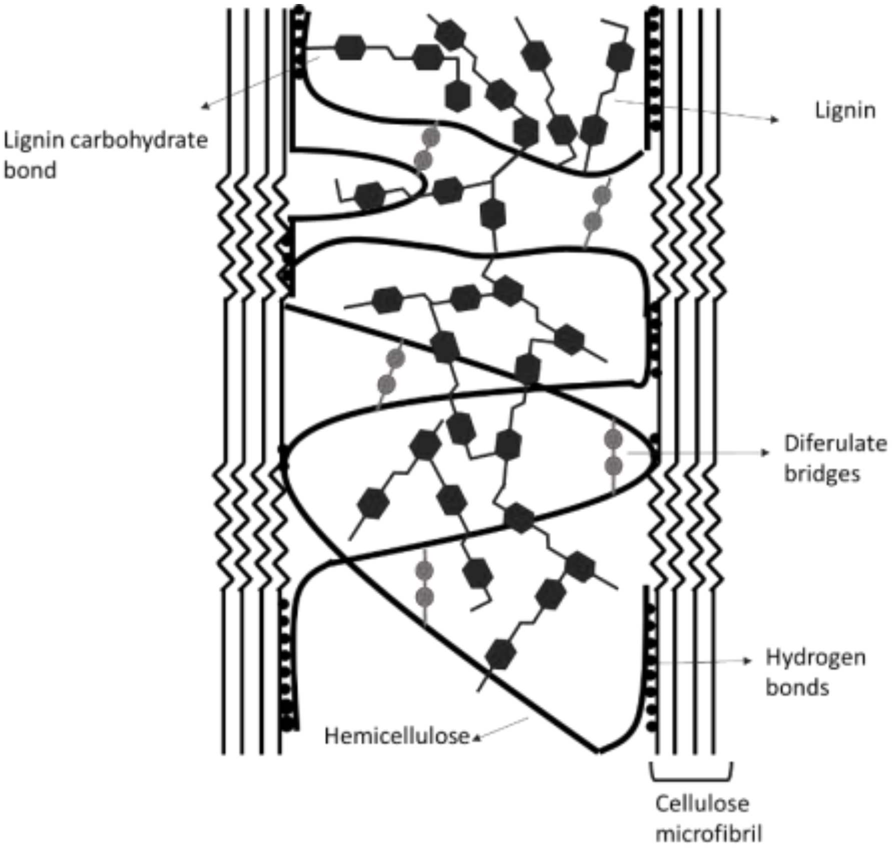


Figure 2.15: Schematic showing the polymer interactions in the secondary cell wall[6, 116, 131]

Interesting, an early study showed changes in orientation of preferential localization of AcGX between the S layers of woody plants using in-situ labeling experiments.
The S1 and S3 layers have different cellulose orientation than in S2 layer. Hence, the presence of AcGX at the transition zone of S layer is suggestive that they play a role in cellulose organization.[135] All these results show strong bonding interaction between cellulose and hemicellulose.

Xylans have side chains which are all present on one side of the polymer with sometimes varying degree of acetylation of sugar monomers.[136-138] Due to the huge variations of side chains it can be very challenging to determine the binding of such xylans.[87] Initial modeling studies using molecular dynamics simulations showed that xylan with no side chains interacts with the hydrophilic phase of the cellulose and with 2-fold screw configuration.[139] This was validated experimentally by performing solid-state NMR of the plant cell wall. Modeling studies of xylan were conducted in which xylan had O-acetylation substituents. It was found that these groups stabilize the binding of the xylan to the cellulose.[140] This was contrary to previous results because it was believed that xylan molecules have to be linear and without substitution to have maximum H-bonds with the cellulose. The results were rationalized by proposing that xylan with no or low substitutions bind to the cellulose, and the less side chain containing xylan interact with other xylan polymers. This causes xylan molecules to aggregate with xylan molecules that are directly bound to cellulose.[141-142] Modeling studies of xylan, which had side chains of arabinose and galacturonic acid, were also conducted. It was found that xylan with such substituents interacts with (110) hydrophilic face of cellulose. α-1,2- substitutions could stabilize the adsorbed xylan by restricting its mobility on cellulose surface.[141] Strong stabilization was also observed when multiple xylan chains with galacturonic acid were present due to calcium forming cross-links between the galacturonic acid of adjacent xylan chains.[141]

Lignin polysaccharide interactions

Lignin and hemicelluloses such as xylan and glucomannans are simultaneously formed and deposited in the secondary cell wall. Lignin is found in two fractions when extracted from the cell wall by using mixture of dioxane: water.[59] One of the fractions contains pure lignin in low yield while the other fraction has hemicellulose and pectin sugar monomers linked to lignin. Additionally, the lignin stream obtained during kraft pulping process also contains covalently linked polysaccharides. Our knowledge of lignin biosynthesis and characterization of extracted lignin give indirect evidence of covalent bond to exist between polysaccharides such as hemicellulose, pectins, and lignin. However, there is no clear evidence to know if such bonds are naturally occurring in the wood or are formed during extractions processes. There are a few mild extraction processes that have been used but the challenge to conclusive tease out the covalent bond by the currently available analytical tools remains challenging.[55]

Nonetheless, based on the functional groups of hemicellulose and lignin, and free radical formation during lignin polymerization, possibility of ester or ether linkages between hemicellulose and lignin cannot be discounted. Figure 2.16 shows the different ways in which these bonds could be formed between these two cell wall polymers.

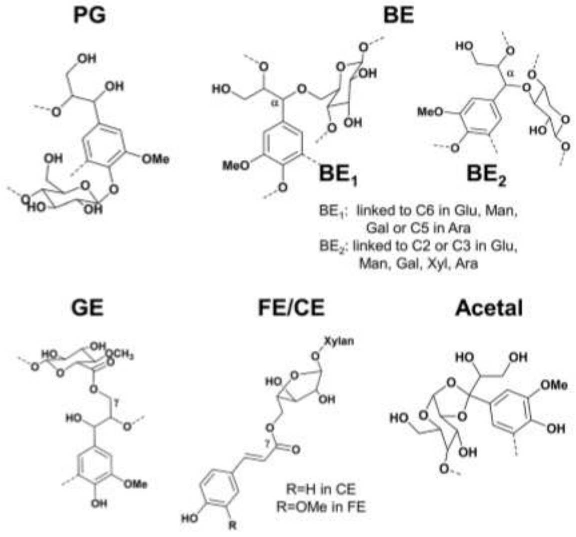


Figure 2.16: Possible ester and ether linkages between cell wall lignin and polysaccharides. PG = phenyl glycosides, BE = benzyl ethers; GE = γ-esters; FE = ferulate esters; CE = coumarate esters.[55]

Ester linkages between terminal nonreducing glycosyl residue and less common monolignols such as p-coumaric acid, ferulic acid, and hydrocinnamic acid have been reported in several different plant species.[87] For instance, ester linkage between ferulic acid and the arabinofuranosyl residue of arabinoxylan and terminal xylosyl residue of xyloglucan in grasses have been identified. Other examples include ester linked p-coumaric acid and arabinoxylan in bamboo, ester linked ferulic acid and terminal arabinosyl and galactopyranosyl residue of pectin present in sugar beet and spinach. Occurrence of biphenyl linkage has also been proposed to form by oxidative reaction between two ferulic acid residues (Figure 2.17). Each of the ferulate, in turn, may be connected to arabinoxylan by ester linkage leading to formation of cross-linked cell wall polymer.[9]

There is no *in vivo* evidence that shows ester and ether linkages between lignin and softwood and hardwood hemicelluloses such as mannan and xylan.[55] Although these hemicelluloses have nucleophilic group and a mechanism for covalent linkages to exist has been proposed. Nucleophilic groups such as hydroxyl or carboxylic acid may re-aromatize the quinone methide intermediate to form ether or ester linkages respectively at the β position of the monolignol.[54] Indirect evidence suggestive of cross-links between carboxylic acid and lignin was found when engineered mutant plant cell wall deficient in galacturonic acid or glucuronic acid were easier to digest by hydrolytic enzymes than wild-type. Similarly, plants in which caffeic acid O-methyl transferase (COMT) gene was downregulated were less recalcitrance to enzyme hydrolysis than wild-type. For COMT plant, an unconventional monolignol, 5-hydroxyl coniferyl alcohol was made that participated in formation of benzodioxane units.[166]

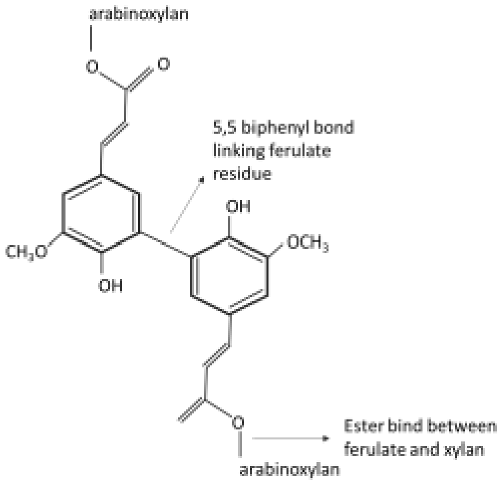


Figure 2.17: Ferulate bridge connecting two ferulic acid molecules at C5 position. The ferulic acid may be connected to arabinoxylan by ester linkage[9]

It was proposed that polysaccharide were not able to couple with lignin having such linkages due to the reduced recalcitrance was observed. Besides covalent linkages electrostatic interaction between xylan and lignin has also been observed. Lignin is hydrophobic and tends to self-aggregates to form particles of nanosize that were found to have extensive surface contact with xylan. Further, using solid-state NMR spectroscopy, electrostatic interactions between lignin and the polar motifs of xylan were reported in intact maize stems.[143]

2.3). Biomass recalcitrance

Due to the intricate arrangement of the cell wall polymers the structure of the cell wall is rigid. This is not favorable as it limits the access of enzymes that degrade cellulose to glucose. Several physical and chemical features of the cell wall that restrict hydrolytic enzyme access are discussed below.

2.3.1). Cellulose accessible surface

Factors such as the particle size of the biomass after milling[144-145] and the porosity[146-148] of the biomass were found to determine the cellulose accessible. Hydrolytic enzymes get absorbed onto the cellulose surface and hydrolyze the cellulose at a solid-liquid interface.[149] It is probable to expect that that greater surface area would result in more absorption or contact of the enzyme with cellulose surface due to which the glucan release would be positively influenced. While this result has been reported by several studies,[144, 146-148] there are some studies that have reported otherwise.[150-151] Specifically, hydrolytic enzyme binds poorly to the surface of cellulose III as compared to cellulose I, yet the glucan release is several folds higher from cellulose III than cellulose I.[150] This result shows that the inherent crystalline structure of cellulose with which the enzyme interacts is important in determining the glucan release. Further, for the same crystalline structure, it was reported that the cellulase enzymes with the carbohydrate binding domain bind specifically to the exposed hydrophobic surface, and it also would influence the mode of action for the enzyme.[151] Nonetheless, while crystal structure of cellulose is important in determining glucan release due to enzymatic hydrolysis, if there is no difference in crystal structure or type of exposed cellulose surface, the cellulose accessible surface would become a determining factor in governing enzymatic hydrolysis of cellulose.

2.3.3). Hemicellulose

Removal of hemicellulose is reported to increase glucan release from cellulose post enzyme hydrolysis. This result suggests that the presence of hemicellulose in lignocellulosic biomass contributes to recalcitrance.[152] Further, hemicellulose is believed to form covalent linkages with lignin.[153-154] It has also been proposed that branched residues of the hemicellulose face away from cellulose and interact with lignin.[155] This type of lignin and hemicellulose arrangement on the cellulose surface restricts the access of hydrolytic enzymes towards cellulose.[156]The sugar residues in hemicellulose have a certain degree of acetylation. The role of acetyl groups on hemicellulose is not completely known, but it has been proposed that acetate groups face towards the lignified region of the cell wall.[140] This is thought to increase the stiffness of the lignin hemicellulose network and make the cell wall rigid.[156] Acetyl groups are also known to affect enzymatic hydrolysis of cellulose. The acetyl groups are converted to acetic acid during dilute acid (commonly sulphuric acid) or hot water pretreatments and the so formed acetic acid is reported to positively[157] and negatively affect glucan yield. As it can reduce pH of the pretreatment liquor, the acetic acid acts as a catalyst and increases the rate of biomass breakdown. Further, pretreatment with acetic acid is known to remove lignin to some extend from biomass and improve enzyme accessibility. However, several studies show that deacetylated biomass is more digestible than biomass containing acetylated hemicellulose.[138, 149, 158]

2.3.4) Lignin

Similar to hemicellulose, removal of lignin is also known to improve cellulose accessibility.[159] Lignin is shown to bind cellulolytic enzymes irreversible, making the enzymes ineffective.[160] Various physical and chemical properties of the lignin structure are reported to influence biomass recalcitrance.[70, 161-162] Physical properties such as lignin content and molecular weight are shown to influence biomass recalcitrance.[70, 163] Chemical properties include the ratio of syringyl (S) to guaciyl (G) monomer composition of lignin.[162] It is shown that S rich lignin has a lower glass transition temperature than G-rich lignin, and it is prone to get re-distributed into lignin aggregates post dilute acid pretreatment faster than G-rich lignin.[164] The higher content of beta aryl ether linkage is found in S-rich lignin, which is linear polymer as compared branched G-rich lignin.[165] The branched lignin is proposed to increase the rigidity of the cell wall and reduce enzymes from accessing cellulose. Unconventional lignin monomer such as caffeyl alchol containing lignin in biomass is also reported to be easier to digest.[166] Lignin residues or derivatives in the pretreated biomass such as presence of vanillin, synringaldehyde, trans-cinnamic acid, hydroxybenzoic acid are shown to inhibit certain cellulase and hemicellulases of the cocktail mix used for hydrolyzing cellulose post-pretreatment.[167]

2.4). Thermochemical pretreatments

Due to the reasons mentioned above, the hydrolytic enzyme access to the cellulose surface is limited. The lignocellulosic biomass is treated with either a particular or a combination of different processes before enzymatic hydrolysis. The primary objective of the pretreatment process is to breakdown the rigid cell wall structure to allow access to hydrolytic enzymes (Figure 2.18). Depending on the type of process applied to disrupt the cell wall, the pretreatment of biomass can be divided into three categories. These include A). Physical disruption with milling B). Use of chemicals- acidic, alkaline, organosolv, or ionic liquids C). Use of enzymes or microbes such as white-rot fungus. The changes to the cell wall structure are dependent upon the type of pretreatment, but changes to the cellulose crystalline structure, reduction of the cellulose DP, (partial) removal of hemicellulose and/or removal of lignin have been reported. One outcome common to all pretreatment processes is increase in cellulose accessibility for hydrolytic enzymes. Due to this at least 5-10-fold increase in glucose yield are obtained post enzymatic hydrolysis of pretreated biomass as compared to native biomass.

2.4.1). Physical pretreatment

Plant biomass is fibrous and one way of increasing cellulose accessibility is by reducing the length of fiber. Some typical methods include vibratory and compression-based ball milling of either dry or wet biomass, chipping, and grinding.[168] The particle size of the biomass is decided based on the next pretreatment step.[169] Generally, the biomass is cut into sizes ranging from meter to centimeter, and depending on the milling process applied sizes of 50 to 500 µm can be obtained. Ultra-fine milling can also be done to obtain particles of size less than 20 µm, but current limitation is of 1um. The composition of cell wall polymers is variable in plants due to which a wide distribution in mechanical properties such as Young's modulus and tensile strength are observed. The stiffness of the material can be known by Young’s modulus while the amount of stress needed to break the material is known by the tensile strength.

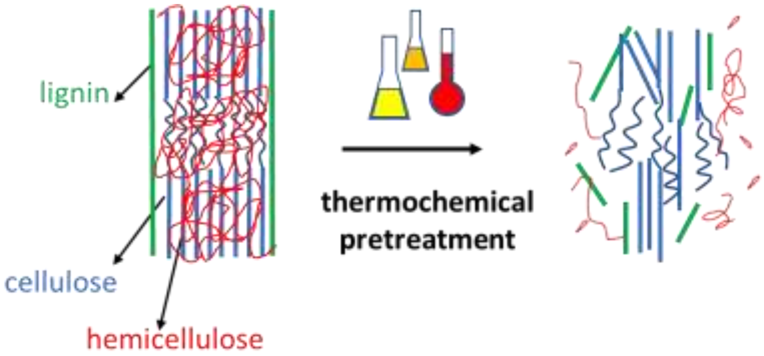


Figure 2.18: Schematic showing cellulose (blue), hemicellulose (red) and lignin (green) in the rigid cell wall which is disrupted due to thermochemical pretreatment

The cost of size reduction depends on the mechanical properties.[168] While ultra-fine milling significantly improves cellulose accessibility, it also increases the overall cost of making biofuel.[170] However, for the purpose storage and handing logistics, physical pretreatment for biomass comminution to few centimeter or millimeter size is necessary.[171]

2.4.2). Chemical pretreatment

The use of chemicals is most effective at making the biomass more conducive for enzymatic hydrolysis. Although before chemical pretreatment, the biomass is often milled using the above-mentioned physical pretreatment.[172] The mode of action of the chemicals usually leads to chemical modification of the cell wall components.[173-174]

Acid pretreatment

The use of different acids ranging from mineral acids (such as sulfuric acid, nitric acid) and organic acids (carbonic, acetic, succinic, fumaric, maleic, citric acid) at variable concentrations (dilute or concentrated) have been for pretreatment of biomass.[175] Water at high temperatures also acts as a weak acid and has been used for pretreating biomass.

Concentrated/Strong Acid

Concentrated sulfuric acid (30-40%) is the most commonly used acid, but use of concentrated hydrochloric, nitric, and trifluoroacetic acid has also been reported.[174] Nikolaus A. Otto and Henry Ford were among the first interested in using sulfuric acid for producing fuels from lignocellulose.[176] The optimal temperature for conducting pretreatment with concentrated sulfuric acid ranges from ambient to 100˚C. As concentrated acid hydrolyzes the polysaccharides like cellulose and hemicellulose to monomeric form, the step of enzymatic hydrolysis is not needed. Additionally, monomeric sugar like glucose and xylose do not degrade into 5-hydroxymethyl furfural and furfural do as the temperature is relatively low during the pretreatment.[177] While there are advantages of using concentrated acid, corrosion of reactor, which in which the pretreatment is conducted occurs over time. This increases the cost along with this, using acid in its concentrated form is also expensive. Since the pH of the pretreatment liquor is low, it has to be neutralized, and an additional step of ion-exchange chromatography is at times done to separate the sugars and the acid. Recovery of acid after the pretreatment process is critical to make the process commercially viable. Methods such as distillation and vacuum evaporation are used for recovering spent acid for reuse. At industrial scale, concentrated acid has been used for pretreating biomass by Arkenol Inc., Masada Resource Group, BlueFire Renewables, Virdia, and Biosulfurol Energy.[176]

Dilute/Weak Acid

Of all the pretreatments, the combination of milling followed by the use of hot water or dilute acid is being used at a commercial scale in biorefineries. The effect of hot water and dilute acid on different lignocellulosic biomass and each of the cell wall components have been widely studied.[178] Dilute acid is most commonly used solvent for pretreatment of lignocellulosic biomass.[179] Acids such as sulfuric acid, hydrochloric acid, nitric acid, peracetic acid or phosphoric acid have been reported to be used at 0.05-5% (w/v) concentrations for pretreatment of various feedstocks. Dilute acid pretreatments are performed at temperatures ranging from 140-220 °C with the pressure generally being less than 10 atmospheres. The type of reactor in which the dilute acid pretreatment is conducted is dependent on the temperature. For temperatures less than 160 °C a flow-through reactor is used and the biomass loading is 5-10% (w/v) while for temperatures greater than 160 °C a stationary batch mode reactor with loading as high as 10-40% is used.[180] The temperature of the biomass filled reactor is increased to the pretreatment temperature and is maintained for a period of time (retention period) at the pretreatment temperature. The period of time and pretreatment temperature defines the pretreatment severity. Usually, for dilute acid pretreatment, if retention period is few minutes then the temperature is above 180 ˚C and if the retention period is few hours then the temperature is below 180 ˚C. The balance between time and temperature is important to maintain for reducing the sugar degradation to inhibitor products such as HMF and furfural.[177-178] The severity of the pretreatment can be described by using the following equations. These equations combine time, temperature and pH of the media to give the severity factor. [181]

$$\log(S) = \log(R) - pH$$

$$R = t \times e^{\left[\frac{(T-100)}{14.75}\right]}$$

Mode of Action

Protons released from the acid catalyze the cleavage of the glycosidic linkage of the polysaccharides such as cellulose and hemicellulose. Xylose and glucose monomers are released to the pretreatment liquor which at high severity of pretreatment is shown to degrade to 5-HMF and furfural. Further conversion of furfural to formic acid and acetic acid has also been reported. Furfural and/or HMF can undergo further rearrangements and produce aromatics. These aromatics are reported to undergo polymerization and/or polycondensation reactions can form pseudo-lignin.[178] The effect of dilute acid leads to structural changes on the cell wall polymers cellulose and lignin. Some studies have also shown that lignin becomes fluid and then coalesce to form droplets within the cell wall matrix.[159, 182] Other studies have proposed lignin to undergo condensation reactions to form lignin aggregates. [165]These lignin aggregates or droplets are clearly visible in scanning electron microscopy images.[159] Changes in the cellulose ultrastructure cellulose have also been reported due to dilute acid pretreatment. These changes include cellulose crystallinity, cellulose form and degree of polymerization. The crystallinity was found to increase and a closer examination by NMR revealed rapid hydrolysis of the more solvent accessible amorphous region followed by much slower hydrolysis of the crystalline cellulose. The loss of amorphous region is reported to the reason for increase in crystallinity.[183] However, at higher temperatures the glucose released from the amorphous part of cellulose is susceptible to further degradation.[177, 184] Dilute acid pretreatment is also known to cause changes in the cellulose ultrastructure. These include due to the presence of water during dilute acid pretreatment change in the relatively low Iα content to Iβ.[29] The crystallite size is also reported to increase due to coalescing of microfibrils.[185] Despite certain drawbacks such as corrosion of pretreatment vessel over time due to acidic conditions, loss of fermentable sugars at high severity conditions and formation of enzyme inhibitory compounds like furfural or 5-hydroxy methyl furfural, dilute acid is still a preferred method for pretreating lignocellulosic biomass.[186]

Alkali pretreatment

Calcium hydroxide and sodium hydroxide

Common chemicals used in this pretreatment include calcium hydroxide (also called lime), potassium hydroxide or sodium hydroxide.[187-188] Pretreatment of lignocellulosic biomass with calcium hydroxide offers certain advantages like low cost due to easy availability, nontoxicity, easy recovery and can be recycled. The mode of action involves the removal of lignin but is slow and that makes the pretreatment commercially expensive.[189] An example of alkali pretreatment of biomass is seen in the paper and pulp industry.[176, 190] Biomass is cooked with a combination of sodium hydroxide and sodium disulfide and sulfur derivates are formed that react with lignin eventually removing it from biomass. Additionally, alkali such as sodium hydroxide is also reported to removes acetyl and uronic acid groups of hemicellulose and pectins.[39, 188] It also swells the lignocellulosic biomass, reduced cellulose crystallinity and hence increases the internal surface area for enzymatic hydrolysis.[187] This pretreatment is performed at lower temperatures (50 to 120 ˚C) but takes a longer period (3 to 4 days) to complete its action. It is usually preferred for biomass that has low lignin content such as herbaceous grasses or softwoods. As the temperature conditions are mild, formation of inhibitory products like HMF and furfural are limited. [191]Addition of oxygen or air is reported to enhance the rate of removal of lignin. As compared to sodium hydroxide, calcium hydroxide is cheaper and has been also proven to show higher cellulose yield.[189]

Ammonia Pretreatment

Ammonia based pretreatments include the ammonia fiber explosion-method (AFEX), ammonia recycle percolation (ARP)[192] and soaking in aqueous ammonia (SAA).[193] Liquid ammonia is added to pre-wet biomass and pretreatment is done 60 to 140 ˚C for 5 to 45 minutes before releasing the ammonia. Ammonia is volatile due to which it can be easily recovered and recycled. Similar to other alkalis, AFEX results in de-acetylation and lignin removal.[194]

A flow-through reactor consisting of a column reactor packed with biomass under pressure ranging from 2-3 MPa with recycling ammonia through percolation is used for ARP pretreatment. The temperature is higher than AFEX and is about 160-180 ˚C with 5-15% ammonia containing aqueous solution. After reaction there are two fractions 1). solid fraction which is rich in cellulose and hemicellulose and 2). a liquid fraction that has ammonia along with lignin and some sugars that mostly coming from hemicellulose.[194]

For soaking in aqueous ammonia (SAA) pretreatment, as the name suggests, the biomass is soaked in aqueous temperature at mild temperatures of 30-75 ˚C but for long periods sometimes ranging for days. Just AFEX and ARP, SAA removes lignin from the biomass, and a solid fraction reach in hemicellulose and cellulose is obtained. For all ammonia-based pretreatments, cost of ammonia is a limiting factor for using it at a commercial and large scale. However, ammonia can be recovered and recycled which makes it an attractive pretreatment option.[193]

Organosolv

In this pretreatment, the lignocellulosic biomass is treated with an organic solvent or mixtures of organic solvents with water. Reported organic solvents include tetrahydrofuran, acetone, glycerol, alcohols such as methanol and ethanol, glycols such as ethylene glycol and triethylene glycol.[69, 73, 195] Sometimes acids such as HCl, H_2SO_4, oxalic acid, acetylsalicylic acid, and salicylic acid are added and these act as catalysts. The pretreatment is done in temperature ranging from 100 to 250 ˚C. The organic solvents act on the lignin and remove it. The efficiency of the pretreatment is dependent on several parameters such as ratio of solvent to water, temperature and solvent to solid loading.[196] The disadvantage of Organosolv based pretreatment is that it could act as an inhibitor during the enzyme hydrolysis and fermentation steps. Due to this partial to complete removal of solvent is necessary and some of the organic solvents should be properly discarded as they are can be harmful to the environment. This pretreatment is preferred when high quality lignin which can be valorized to bioproducts is to be extracted for lignocellulosic biomass. If the solvents are removed from the pretreated material then it has been reported that absorption of cellulase enzymes is high and there is increased cellulose accessible surface due to removal of lignin.[190]

Ionic liquids (IL)-based pretreatment

Ionic liquids are salts made of inorganic anions and organic cations. At room temperature, they are in the liquid phase due to the weak interactions of the ions. The ILs work in a different way than the rest of the pretreatments. The lignocellulosic biomass is added to ILs and components such lignin, cellulose and hemicellulose get solubilized in the ILs.[197] The insoluble residue is removed and the filtrate is precipitated with antisolvents such as water, methanol or ethanol to recover the cellulose. Due to the difference in polarity ILs, can be selective solvents of lignin or cellulose separation and extraction. Solvent that contains anions of chloride, phosphonate, formate, acetate or alkyl could form hydrogen bonds with cellulose and could dissolve cellulose. While ILs are very effective in separating cell wall components and can be reused, still the ILs are very expensive and not much is known about its combability with enzymes.[179, 186, 198]

2.4.3). Biological pretreatment

Microorganisms such as white, brown and soft rot-fungi could degrade hemicellulose and lignin. The C-C lignin backbone is cleaved by white-rot microbes secreted enzymes such as lignin peroxidases. These enzymes work in the presence of hydrogen peroxide. The hemicelluloses are hydrolyzed by different types of glycosyl hydrolases.[199] Laccases and superoxide dismutase have also been added to biomass as they aerobically catalyzed lignin degradation.[200] Biological pretreatments as compared to chemical and physical pretreatments need low energy and have milder conditions. However, the rate of degradation is usually very slow due to which longer residence times are needed, and turnaround is slow.[199]

2.5). References

1. Burton, R. A.; Gidley, M. J.; Fincher, G. B., Heterogeneity in the chemistry, structure and function of plant cell walls. *Nature Chemical Biology* **2010,** *6* (10), 724-732.
2. Li, Y.; Horsman, M.; Wu, N.; Lan, C. Q.; Dubois-Calero, N., Biofuels from microalgae. *Biotechnology progress* **2008,** *24* (4), 815-820.
3. Hess, J. R.; Wright, C. T.; Kenney, K. L., Cellulosic biomass feedstocks and logistics for ethanol production. *Biofuels, Bioproducts and Biorefining* **2007,** *1* (3), 181-190.
4. Ragauskas, A. J.; Williams, C. K.; Davison, B. H.; Britovsek, G.; Cairney, J.; Eckert, C. A.; Frederick, W. J.; Hallett, J. P.; Leak, D. J.; Liotta, C. L.; Mielenz, J. R.; Murphy, R.; Templer, R.; Tschaplinski, T., The Path Forward for Biofuels and Biomaterials. *Science* **2006,** *311* (5760), 484.
5. Plomion, C.; Leprovost, G.; Stokes, A., Wood Formation in Trees. *Plant Physiology* **2001,** *127* (4), 1513.
6. Mellerowicz, E. J.; Baucher M Fau - Sundberg, B.; Sundberg B Fau - Boerjan, W.; Boerjan, W., Unravelling cell wall formation in the woody dicot stem. (0167-4412 (Print)).
7. Cosgrove, D. C., Comparative structure and biomechanics of plant primary and secondary cell walls. *Frontiers in plant science* **2012,** *3*, 204.
8. Keegstra, K., Plant Cell Walls. *Plant Physiology* **2010,** *154* (2), 483.
9. Keegstra, K.; Talmadge, K. W.; Bauer, W. D.; Albersheim, P., The Structure of Plant Cell Walls. *Plant Physiology* **1973,** *51* (1), 188.
10. Cox, E. G.; Jeffrey, G. A., Crystal Structure of Glucosamine Hydrobromide. *Nature* **1939,** *143* (3630), 894-895.
11. Nomenclature of Carbohydrates, (Recommendations 1996). In *Advances in Carbohydrate Chemistry and Biochemistry*, Horton, D., Ed. Academic Press: 1997; Vol. 52, pp 44-177.
12. O'Sullivan, A. C., Cellulose: the structure slowly unravels. *Cellulose* **1997,** *4* (3), 173-207.
13. Gardner, K. H.; Blackwell, J., The hydrogen bonding in native cellulose. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1974,** *343* (1), 232-237.
14. Delmer, D. P., Cellulose Biosynthesis. *Annual Review of Plant Physiology* **1987,** *38* (1), 259-290.
15. Tarchevskiĭ, I. A., *Cellulose: Biosynthesis and Structure*. Berlin, Heidelberg : Springer Berlin Heidelberg: 1991.
16. Zimmer, J., A Molecular Description of Cellulose Biosynthesis. *Biophysical Journal* **2015,** *108* (2), 499a-499a.
17. Nishiyama, Y.; Langan, P.; Chanzy, H., Crystal Structure and Hydrogen-Bonding System in Cellulose Iβ from Synchrotron X-ray and Neutron Fiber Diffraction. *Journal of the American Chemical Society* **2002,** *124* (31), 9074-9082.
18. Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P., Crystal Structure and Hydrogen Bonding System in Cellulose Iα from Synchrotron X-ray and Neutron Fiber Diffraction. *Journal of the American Chemical Society* **2003,** *125* (47), 14300-14306.
19. Funahashi, R.; Okita, Y.; Hondo, H.; Zhao, M.; Saito, T.; Isogai, A., Different Conformations of Surface Cellulose Molecules in Native Cellulose Microfibrils Revealed by Layer-by-Layer Peeling. *Biomacromolecules* **2017,** *18* (11), 3687-3694.
20. Ding, S.-Y.; Himmel, M. E., The Maize Primary Cell Wall Microfibril: A New Model Derived from Direct Visualization. *Journal of Agricultural and Food Chemistry* **2006,** *54* (3), 597-606.

21. Cosgrove, D. J., Re-constructing our models of cellulose and primary cell wall assembly. *Current Opinion in Plant Biology* **2014,** *22*, 122-131.

22. Kubicki, J. D.; Yang, H.; Sawada, D.; O'Neill, H.; Oehme, D.; Cosgrove, D., The Shape of Native Plant Cellulose Microfibrils. *Scientific Reports* **2018,** *8* (1), 13983.

23. Vandavasi, V. G.; Putnam, D. K.; Zhang, Q.; Petridis, L.; Heller, W. T.; Nixon, B. T.; Haigler, C. H.; Kalluri, U.; Coates, L.; Langan, P.; Smith, J. C.; Meiler, J.; O'Neill, H., A Structural Study of CESA1 Catalytic Domain of Arabidopsis Cellulose Synthesis Complex: Evidence for CESA Trimers. *Plant Physiology* **2016,** *170* (1), 123.

24. Newman, R. H.; Hill, S. J.; Harris, P. J., Wide-Angle X-Ray Scattering and Solid-State Nuclear Magnetic Resonance Data Combined to Test Models for Cellulose Microfibrils in Mung Bean Cell Walls. *Plant Physiology* **2013,** *163* (4), 1558.

25. Fernandes, A. N.; Thomas, L. H.; Altaner, C. M.; Callow, P.; Forsyth, V. T.; Apperley, D. C.; Kennedy, C. J.; Jarvis, M. C., Nanostructure of cellulose microfibrils in spruce wood. *Proceedings of the National Academy of Sciences* **2011,** *108* (47), E1195.

26. Oehme, D. P.; Doblin, M. S.; Wagner, J.; Bacic, A.; Downton, M. T.; Gidley, M. J., Gaining insight into cell wall cellulose macrofibril organisation by simulating microfibril adsorption. *Cellulose* **2015,** *22* (6), 3501-3520.

27. Isogai, A.; Usuda, M.; Kato, T.; Uryu, T.; Atalla, R. H., Solid-state CP/MAS carbon-13 NMR study of cellulose polymorphs. *Macromolecules* **1989,** *22* (7), 3168-3172.

28. Atalla Rh Fau - Vanderhart, D. L.; Vanderhart, D. L., Native cellulose: a composite of two distinct crystalline forms. (0036-8075 (Print)).

29. Wada, M.; Sugiyama, J.; Okano, T., Native celluloses on the basis of two crystalline phase (Iα/Iβ) system. *Journal of Applied Polymer Science* **1993,** *49* (8), 1491-1496.

30. French, A. D.; Santiago Cintrón, M., Cellulose polymorphy, crystallite size, and the Segal Crystallinity Index. *Cellulose* **2013,** *20* (1), 583-588.

31. Park, S.; Baker, J. O.; Himmel, M. E.; Parilla, P. A.; Johnson, D. K., Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnology for Biofuels* **2010,** *3* (1), 10.

32. Åkerholm, M.; Hinterstoisser, B.; Salmén, L., Characterization of the crystalline structure of cellulose using static and dynamic FT-IR spectroscopy. *Carbohydrate Research* **2004,** *339* (3), 569-578.

33. Ruland, W., X-ray determination of crystallinity and diffuse disorder scattering. *Acta Crystallographica* **1961,** *14* (11), 1180-1185.

34. Garvey, C. J.; Parker, I. H.; Simon, G. P., On the Interpretation of X-Ray Diffraction Powder Patterns in Terms of the Nanostructure of Cellulose I Fibres. *Macromolecular Chemistry and Physics* **2005,** *206* (15), 1568-1575.

35. Atalla, R. H.; VanderHart, D. L., The role of solid state 13C NMR spectroscopy in studies of the nature of native celluloses. *Solid State Nuclear Magnetic Resonance* **1999,** *15* (1), 1-19.

36. Newman, R. H., Homogeneity in cellulose crystallinity between samples of Pinus radiata wood. In *Holzforschung*, 2004; Vol. 58, p 91.

37. Thygesen, A.; Oddershede, J.; Lilholt, H.; Thomsen, A. B.; Ståhl, K., On the determination of crystallinity and cellulose content in plant fibres. *Cellulose* **2005,** *12* (6), 563.

38. Larsson, P. T.; Hult, E.-L.; Wickholm, K.; Pettersson, E.; Iversen, T., CP/MAS 13C-NMR spectroscopy applied to structure and interaction studies on cellulose I. *Solid State Nuclear Magnetic Resonance* **1999,** *15* (1), 31-40.

39. Kumar, R.; Mago, G.; Balan, V.; Wyman, C. E., Physical and chemical characterizations of corn stover and poplar solids resulting from leading pretreatment technologies. *Bioresource Technology* **2009,** *100* (17), 3948-3962.

40. Hallac, B. B.; Ragauskas, A. J., Analyzing cellulose degree of polymerization and its relevancy to cellulosic ethanol. *Biofuels, Bioproducts and Biorefining* **2011,** *5* (2), 215-225.

41. Dupont, A.-L.; Mortha, G., Comparative evaluation of size-exclusion chromatography and viscometry for the characterisation of cellulose. *Journal of Chromatography A* **2004,** *1026* (1), 129-141.

42. Sjöholm, E., Size exclusion chromatography of cellulose and cellulose derivatives. *Handbook Of Size Exclusion Chromatography And Related Techniques: Revised And Expanded* **2003,** *91*.

43. French, A. D., Idealized powder diffraction patterns for cellulose polymorphs. *Cellulose* **2014,** *21* (2), 885-896.

44. Nishiyama, Y.; Kuga, S.; Okano, T., Mechanism of mercerization revealed by X-ray diffraction. *Journal of Wood Science* **2000,** *46* (6), 452-457.

45. Langan, P.; Nishiyama, Y.; Chanzy, H., A Revised Structure and Hydrogen-Bonding System in Cellulose II from a Neutron Fiber Diffraction Analysis. *Journal of the American Chemical Society* **1999,** *121* (43), 9940-9946.

46. Porro, F.; Bédué, O.; Chanzy, H.; Heux, L., Solid-State 13C NMR Study of Na−Cellulose Complexes. *Biomacromolecules* **2007,** *8* (8), 2586-2593.

47. Wada, M.; Chanzy, H.; Nishiyama, Y.; Langan, P., Cellulose IIII Crystal Structure and Hydrogen Bonding by Synchrotron X-ray and Neutron Fiber Diffraction. *Macromolecules* **2004,** *37* (23), 8548-8555.

48. Nishimura, H.; Sarko, A., Mercerization of cellulose. 6. Crystal and molecular structure of Na-cellulose IV. *Macromolecules* **1991,** *24* (3), 771-778.

49. Gardiner, E. S.; Sarko, A., Packing analysis of carbohydrates and polysaccharides. 16. The crystal structures of celluloses IVI and IVII. *Canadian Journal of Chemistry* **1985,** *63* (1), 173-180.

50. Zeronian, S. H.; Ryu, H.-S., Properties of cotton fibers containing the cellulose IV crystal structure. *Journal of Applied Polymer Science* **1987,** *33* (7), 2587-2604.

51. Boerjan, W.; Ralph, J.; Baucher, M., Lignin Biosynthesis. *Annual Review of Plant Biology* **2003,** *54* (1), 519-546.

52. Ralph, J.; Lundquist, K.; Brunow, G.; Lu, F.; Kim, H.; Schatz, P. F.; Marita, J. M.; Hatfield, R. D.; Ralph, S. A.; Christensen, J. H.; Boerjan, W., Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenyl- propanoids. *Phytochemistry Reviews* **2004,** *3* (1), 29-60.

53. Boudet, A.-M., Lignins and lignification: selected issues. *Plant Physiology and Biochemistry* **2000,** *38* (1-2), 81-96.

54. Vanholme, R.; Demedts, B.; Morreel, K.; Ralph, J.; Boerjan, W., Lignin Biosynthesis and Structure. *Plant Physiology* **2010,** *153* (3), 895.

55. Giummarella, N.; Pu, Y.; Ragauskas, A. J.; Lawoko, M., A critical review on the analysis of lignin carbohydrate bonds. *Green Chemistry* **2019,** *21* (7), 1573-1595.

56. Grabber, J. H.; Ralph, J.; Hatfield, R. D.; Quideau, S.; Kuster, T.; Pell, A. N., Dehydrogenation Polymer−Cell Wall Complexes as a Model for Lignified Grass Walls. *Journal of Agricultural and Food Chemistry* **1996,** *44* (6), 1453-1459.

57. Lapierre, C.; Pollet, B.; Rolando, C., New insights into the molecular architecture of hardwood lignins by chemical degradative methods. *Research on Chemical Intermediates* **1995,** *21* (3), 397.

58. Gellerstedt, G., Chemical Degradation Methods: Permanganate Oxidation. In *Methods in Lignin Chemistry*, Lin, S. Y.; Dence, C. W., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 1992; pp 322-333.

59. BjÖRkman, A., Isolation of Lignin from Finely Divided Wood with Neutral Solvents. *Nature* **1954,** *174* (4440), 1057-1058.

60. Ralph, J.; Akiyama, T.; Coleman, H. D.; Mansfield, S. D., Effects on Lignin Structure of Coumarate 3-Hydroxylase Downregulation in Poplar. *BioEnergy Research* **2012,** *5* (4), 1009-1019.

61. Lundquist, K., Proton (1H) NMR Spectroscopy. In *Methods in Lignin Chemistry*, Lin, S. Y.; Dence, C. W., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 1992; pp 242-249.

62. Heikkinen, S.; Toikka, M. M.; Karhunen, P. T.; Kilpeläinen, I. A., Quantitative 2D HSQC (Q-HSQC) via Suppression of J-Dependence of Polarization Transfer in NMR Spectroscopy: Application to Wood Lignin. *Journal of the American Chemical Society* **2003,** *125* (14), 4362-4367.

63. Xia, Z.; Akim, L. G.; Argyropoulos, D. S., Quantitative 13C NMR Analysis of Lignins with Internal Standards. *Journal of Agricultural and Food Chemistry* **2001,** *49* (8), 3573-3578.

64. Qu, Y.; Luo, H.; Li, H.; Xu, J., Comparison on structural modification of industrial lignin by wet ball milling and ionic liquid pretreatment. *Biotechnology Reports* **2015,** *6*, 1-7.

65. Foston, M.; Hubbell, C. A.; Samuel, R.; Jung, S.; Fan, H.; Ding, S.-Y.; Zeng, Y.; Jawdy, S.; Davis, M.; Sykes, R.; Gjersing, E.; Tuskan, G. A.; Kalluri, U.; Ragauskas, A. J., Chemical, ultrastructural and supramolecular analysis of tension wood in Populus tremula x alba as a model substrate for reduced recalcitrance. *Energy & Environmental Science* **2011,** *4* (12), 4962-4971.

66. Liu, Y.; Hu, T.; Wu, Z.; Zeng, G.; Huang, D.; Shen, Y.; He, X.; Lai, M.; He, Y., Study on biodegradation process of lignin by FTIR and DSC. *Environmental Science and Pollution Research* **2014,** *21* (24), 14004-14013.

67. Ragauskas, A. J.; Beckham, G. T.; Biddy, M. J.; Chandra, R.; Chen, F.; Davis, M. F.; Davison, B. H.; Dixon, R. A.; Gilna, P.; Keller, M.; Langan, P.; Naskar, A. K.; Saddler, J. N.; Tschaplinski, T. J.; Tuskan, G. A.; Wyman, C. E., Lignin Valorization: Improving Lignin Processing in the Biorefinery. *Science* **2014,** *344* (6185), 1246843.

68. Baumberger, S.; Dole, P.; Lapierre, C., Using Transgenic Poplars to Elucidate the Relationship between the Structure and the Thermal Properties of Lignins. *Journal of Agricultural and Food Chemistry* **2002,** *50* (8), 2450-2453.

69. El Hage, R.; Brosse, N.; Chrusciel, L.; Sanchez, C.; Sannigrahi, P.; Ragauskas, A., Characterization of milled wood lignin and ethanol organosolv lignin from miscanthus. *Polymer Degradation and Stability* **2009,** *94* (10), 1632-1638.

70. Tolbert, A.; Akinosho, H.; Khunsupat, R.; Naskar, A. K.; Ragauskas, A. J., Characterization and analysis of the molecular weight of lignin for biorefining studies. *Biofuels, Bioproducts and Biorefining* **2014,** *8* (6), 836-856.

71. SjÖStrÖM, E., Chapter 4 - LIGNIN. In *Wood Chemistry (Second Edition)*, SjÖStrÖM, E., Ed. Academic Press: San Diego, 1993; pp 71-89.

72. Lupoi, J. S.; Singh, S.; Parthasarathi, R.; Simmons, B. A.; Henry, R. J., Recent innovations in analytical methods for the qualitative and quantitative assessment of lignin. *Renewable and Sustainable Energy Reviews* **2015,** *49*, 871-906.

73. Ramires, E. C.; Megiatto Jr, J. D.; Gardrat, C.; Castellan, A.; Frollini, E., Valorization of an industrial organosolv–sugarcane bagasse lignin: Characterization and use as a matrix in biobased composites reinforced with sisal fibers. *Biotechnology and Bioengineering* **2010,** *107* (4), 612-621.

74. Li, M.; Pu, Y.; Ragauskas, A. J., Current Understanding of the Correlation of Lignin Structure with Biomass Recalcitrance. *Front Chem* **2016,** *4*, 45-45.

75. Scheller, H. V.; Ulvskov, P., Hemicelluloses. *Annual Review of Plant Biology* **2010,** *61* (1), 263-289.

76. Ebringerová, A.; Hromádková, Z.; Heinze, T., Hemicellulose. In *Polysaccharides I: Structure, Characterization and Use*, Heinze, T., Ed. Springer Berlin Heidelberg: Berlin, Heidelberg, 2005; pp 1-67.

77. Vogel, J., Unique aspects of the grass cell wall. *Current Opinion in Plant Biology* **2008,** *11* (3), 301-307.

78. Ebringerová, A.; Heinze, T., Xylan and xylan derivatives – biopolymers with valuable properties, 1. Naturally occurring xylans structures, isolation procedures and properties. *Macromolecular Rapid Communications* **2000,** *21* (9), 542-556.

79. Moreira, L. R. S.; Filho, E. X. F., An overview of mannan structure and mannan-degrading enzyme systems. *Applied Microbiology and Biotechnology* **2008,** *79* (2), 165.

80. Capek, P.; Kubačková, M.; Alföldi, J.; Bilisics, L.; Lišková, D.; Kákoniová, D., Galactoglucomannan from the secondary cell wall of Picea abies L. Karst. *Carbohydrate Research* **2000,** *329* (3), 635-645.

81. Hayashi, T., Xyloglucans in the Primary Cell Wall. *Annual Review of Plant Physiology and Plant Molecular Biology* **1989,** *40* (1), 139-168.

82. Hisamatsu, M.; York, W. S.; Darvill, A. G.; Albersheim, P., Characterization of seven xyloglucan oligosaccharides containing from seventeen to twenty glycosyl residues. *Carbohydrate Research* **1992,** *227*, 45-71.

83. York, W. S.; Impallomeni, G.; Hisamatsu, M.; Albersheim, P.; Darvill, A. G., Eleven newly characterized xyloglucan oligoglycosyl alditols: the specific effects of sidechain structure and location on 1H NMR chemical shifts. *Carbohydrate Research* **1995,** *267* (1), 79-104.

84. Vincken, J. P.; York, W. S.; Beldman, G.; Voragen, A. G. J., Two General Branching Patterns of Xyloglucan, XXXG and XXGG. *Plant Physiology* **1997,** *114* (1), 9.

85. Fry, S. C.; York, W. S.; Albersheim, P.; Darvill, A.; Hayashi, T.; Joseleau, J.-P.; Kato, Y.; Lorences, E. P.; Maclachlan, G. A.; McNeil, M.; Mort, A. J.; Grant Reid, J. S.; Seitz, H. U.; Selvendran, R. R.; Voragen, A. G. J.; White, A. R., An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiologia Plantarum* **1993,** *89* (1), 1-3.

86. Mazumder, K.; Peña, M. J.; O'Neill, M. A.; York, W. S., Structural Characterization of the Heteroxylans from Poplar and Switchgrass. In *Biomass Conversion: Methods and Protocols*, Himmel, M. E., Ed. Humana Press: Totowa, NJ, 2012; pp 215-228.

87. Smith, P. J.; Wang, H.-T.; York, W. S.; Peña, M. J.; Urbanowicz, B. R., Designer biomass for next-generation biorefineries: leveraging recent insights into xylan structure and biosynthesis. *Biotechnology for Biofuels* **2017,** *10* (1), 286.

88. Jacobs, A.; Larsson, P. T.; Dahlman, O., Distribution of Uronic Acids in Xylans from Various Species of Soft- and Hardwood As Determined by MALDI Mass Spectrometry. *Biomacromolecules* **2001,** *2* (3), 979-990.

89. Chong, S.-L.; Virkki, L.; Maaheimo, H.; Juvonen, M.; Derba-Maceluch, M.; Koutaniemi, S.; Roach, M.; Sundberg, B.; Tuomainen, P.; Mellerowicz, E. J.; Tenkanen, M., O-Acetylation of

glucuronoxylan in Arabidopsis thaliana wild type and its change in xylan biosynthesis mutants. *Glycobiology* **2014,** *24* (6), 494-506.

90. Andersson, S.-I.; Samuelson, O.; Ishihara, M.; Shimizu, K., Structure of the reducing end-groups in spruce xylan. *Carbohydrate Research* **1983,** *111* (2), 283-288.

91. Ishii, T., Acetylation at O-2 of arabinofuranose residues in feruloylated arabinoxylan from bamboo shoot cell-walls. *Phytochemistry* **1991,** *30* (7), 2317-2320.

92. Goubet, F.; Barton, C. J.; Mortimer, J. C.; Yu, X.; Zhang, Z.; Miles, G. P.; Richens, J.; Liepman, A. H.; Seffen, K.; Dupree, P., Cell wall glucomannan in Arabidopsis is synthesised by CSLA glycosyltransferases, and influences the progression of embryogenesis. *The Plant Journal* **2009,** *60* (3), 527-538.

93. Bento, J. F.; Mazzaro, I.; de Almeida Silva, L. M.; de Azevedo Moreira, R.; Ferreira, M. L. C.; Reicher, F.; de Oliveira Petkowicz, C. L., Diverse patterns of cell wall mannan/galactomannan occurrence in seeds of the Leguminosae. *Carbohydrate Polymers* **2013,** *92* (1), 192-199.

94. Thakur, B. R.; Singh, R. K.; Handa, A. K.; Rao, M. A., Chemistry and uses of pectin — A review. *Critical Reviews in Food Science and Nutrition* **1997,** *37* (1), 47-73.

95. Mort, A. J.; Qiu, F.; Maness, N. O., Determination of the pattern of methyl esterification in pectin. Distribution of contiguous nonesterified residues. *Carbohydrate Research* **1993,** *247*, 21-35.

96. Willats, W. G. T.; Orfila, C.; Limberg, G.; Buchholt, H. C.; van Alebeek, G.-J. W. M.; Voragen, A. G. J.; Marcus, S. E.; Christensen, T. M. I. E.; Mikkelsen, J. D.; Murray, B. S.; Knox, J. P., Modulation of the Degree and Pattern of Methyl-esterification of Pectic Homogalacturonan in Plant Cell Walls: implications for pectin methyl esterase action, matrix properties, and cell adhesion. *Journal of Biological Chemistry* **2001,** *276* (22), 19404-19413.

97. Ridley, B. L.; O'Neill, M. A.; Mohnen, D., Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* **2001,** *57* (6), 929-967.

98. Gawkowska, D.; Cybulska, J.; Zdunek, A., Structure-Related Gelling of Pectins and Linking with Other Natural Compounds: A Review. *Polymers (Basel)* **2018,** *10* (7), 762.

99. Morris, E. R.; Powell, D. A.; Gidley, M. J.; Rees, D. A., Conformations and interactions of pectins: I. Polymorphism between gel and solid states of calcium polygalacturonate. *Journal of Molecular Biology* **1982,** *155* (4), 507-516.

100. Lau, J. M.; McNeil, M.; Darvill, A. G.; Albersheim, P., Structure of the backbone of rhamnogalacturonan I, a pectic polysaccharide in the primary cell walls of plants. *Carbohydrate Research* **1985,** *137*, 111-125.

101. Komalavilas, P.; Mort, A. J., The acetylation of O-3 of galacturonic acid in the rhamnose-rich portion of pectins. *Carbohydrate Research* **1989,** *189*, 261-272.

102. An, J.; O'Neill Ma Fau - Albersheim, P.; Albersheim P Fau - Darvill, A. G.; Darvill, A. G., Isolation and structural characterization of beta-D-glucosyluronic acid and 4-O-methyl beta-D-glucosyluronic acid-containing oligosaccharides from the cell-wall pectic polysaccharide, rhamnogalacturonan I. (0008-6215 (Print)).

103. Albersheim, P.; Darvill, A.; Roberts, K.; Sederoff, R.; Staehelin, A., *Plant Cell Walls: From Chemistry to Biology*. Garland Science: 2011.

104. Mohnen, D., Pectin structure and biosynthesis. *Current Opinion in Plant Biology* **2008,** *11* (3), 266-277.

105. Darvill, A. G.; McNeil, M.; Albersheim, P., Structure of Plant Cell Walls. *Plant Physiology* **1978,** *62* (3), 418.

106. Shin, K.-S.; Kiyohara, H.; Matsumoto, T.; Yamada, H., Rhamnogalacturonan II dimers cross-linked by borate diesters from the leaves of Panax ginseng C.A. Meyer are responsible for expression of their IL-6 production enhancing activities. *Carbohydrate Research* **1998,** *307* (1), 97-106.

107. O'Neill, M. A.; Ishii, T.; Albersheim, P.; Darvill, A. G., RHAMNOGALACTURONAN II: Structure and Function of a Borate Cross-Linked Cell Wall Pectic Polysaccharide. *Annual Review of Plant Biology* **2004,** *55* (1), 109-139.

108. Somerville, C.; Bauer, S.; Brininstool, G.; Facette, M.; Hamann, T.; Milne, J.; Osborne, E.; Paredez, A.; Persson, S.; Raab, T.; Vorwerk, S.; Youngs, H., Toward a Systems Approach to Understanding Plant Cell Walls. *Science* **2004,** *306* (5705), 2206.

109. McCann, M.; Rose, J., Blueprints for Building Plant Cell Walls. *Plant Physiology* **2010,** *153* (2), 365.

110. Li, X.; Chapple, C., Understanding Lignification: Challenges Beyond Monolignol Biosynthesis. *Plant Physiology* **2010,** *154* (2), 449.

111. Guerriero, G.; Fugelstad, J.; Bulone, V., What Do We Really Know about Cellulose Biosynthesis in Higher Plants? *Journal of Integrative Plant Biology* **2010,** *52* (2), 161-175.

112. Pattathil, S.; Avci, U.; Baldwin, D.; Swennes, A. G.; McGill, J. A.; Popper, Z.; Bootten, T.; Albert, A.; Davis, R. H.; Chennareddy, C.; Dong, R.; Shea, B.; Rossi, R.; Leoff, C.; Freshour, G.; Narra, R.; Neil, M.; York, W. S.; Hahn, M. G., A Comprehensive Toolkit of Plant Cell Wall Glycan-Directed Monoclonal Antibodies. *Plant Physiology* **2010,** *153* (2), 514.

113. Terashima, N.; Fukushima, K., Heterogeneity in formation of lignin—XI: An autoradiographic study of the heterogeneous formation and structure of pine lignin. *Wood Science and Technology* **1988,** *22* (3), 259-270.

114. Voiniciuc, C.; Pauly, M.; Usadel, B., Monitoring Polysaccharide Dynamics in the Plant Cell Wall. *Plant physiology* **2018,** *176* (4), 2590-2600.

115. Pauly, M.; Gille, S.; Liu, L.; Mansoori, N.; de Souza, A.; Schultink, A.; Xiong, G., Hemicellulose biosynthesis. *Planta* **2013,** *238* (4), 627-642.

116. Zhong, R.; Cui, D.; Ye, Z.-H., Secondary cell wall biosynthesis. *New Phytologist* **2019,** *221* (4), 1703-1723.

117. Cosgrove, D. J., Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology* **2005,** *6* (11), 850-861.

118. York, W. S.; Darvill, A. G.; McNeil, M.; Stevenson, T. T.; Albersheim, P., Isolation and characterization of plant cell walls and cell wall components. In *Methods in Enzymology*, Academic Press: 1986; Vol. 118, pp 3-40.

119. McCann, M. C.; Wells, B.; Roberts, K., Direct visualization of cross-links in the primary plant cell wall. *Journal of Cell Science* **1990,** *96* (2), 323.

120. Talbott, L. D.; Ray, P. M., Molecular size and separability features of pea cell wall polysaccharides : implications for models of primary wall structure. (0032-0889 (Print)).

121. Carpita, N. C.; Gibeaut, D. M., Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. (0960-7412 (Print)).

122. Park, Y. B.; Cosgrove, D. J., A Revised Architecture of Primary Cell Walls Based on Biomechanical Changes Induced by Substrate-Specific Endoglucanases. *Plant Physiology* **2012,** *158* (4), 1933.

123. Hayashi, T.; Maclachlan, G., Pea Xyloglucan and Cellulose. *Plant Physiology* **1984,** *75* (3), 596.

124. Zykwinska, A.; Thibault, J.-F.; Ralet, M.-C., Competitive binding of pectin and xyloglucan with primary cell wall cellulose. *Carbohydrate Polymers* **2008,** *74* (4), 957-961.

125. Saladié, M.; Rose, J. K. C.; Cosgrove, D. J.; Catalá, C., Characterization of a new xyloglucan endotransglucosylase/hydrolase (XTH) from ripening tomato fruit and implications for the diverse modes of enzymic action. *The Plant Journal* **2006,** *47* (2), 282-295.

126. Dick-Pérez, M.; Zhang, Y.; Hayes, J.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and Interactions of Plant Cell-Wall Polysaccharides by Two- and Three-Dimensional Magic-Angle-Spinning Solid-State NMR. *Biochemistry* **2011,** *50* (6), 989-1000.

127. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-Pectin Spatial Contacts Are Inherent to Never-Dried Arabidopsis Primary Cell Walls: Evidence from Solid-State Nuclear Magnetic Resonance. *Plant Physiology* **2015,** *168* (3), 871.

128. Pakarinen, A.; Zhang, J.; Brock, T.; Maijala, P.; Viikari, L., Enzymatic accessibility of fiber hemp is enhanced by enzymatic or chemical removal of pectin. *Bioresource Technology* **2012,** *107*, 275-281.

129. Lopez-Sanchez, P.; Martinez-Sanz, M.; Bonilla, M. R.; Wang, D.; Gilbert, E. P.; Stokes, J. R.; Gidley, M. J., Cellulose-pectin composite hydrogels: Intermolecular interactions and material properties depend on order of assembly. *Carbohydrate Polymers* **2017,** *162*, 71-81.

130. DeMartini, J. D.; Pattathil, S.; Avci, U.; Szekalski, K.; Mazumder, K.; Hahn, M. G.; Wyman, C. E., Application of monoclonal antibodies to investigate plant cell wall deconstruction for biofuels production. *Energy & Environmental Science* **2011,** *4* (10), 4332-4339.

131. Mellerowicz, E. J.; Sundberg, B., Wood cell walls: biosynthesis, developmental dynamics and their implications for wood properties. *Current Opinion in Plant Biology* **2008,** *11* (3), 293-300.

132. Barnett, J. R.; Bonham, V. A., Cellulose microfibril angle in the cell wall of wood fibres. *Biological Reviews* **2004,** *79* (2), 461-472.

133. Joseleau, J.-P.; Chambat, G.; Chumpitazi-Hermoza, B., Solubilization of cellulose and other plant structural polysaccharides in 4-methylmorpholine N-oxide: an improved method for the study of cell-wall constituents. *Carbohydrate Research* **1981,** *90* (2), 339-344.

134. Shah, R.; Huang, S.; Pingali, S. V.; Sawada, D.; Pu, Y.; Rodriguez, M.; Ragauskas, A. J.; Kim, S. H.; Evans, B. R.; Davison, B. H.; O'Neill, H., Hemicellulose–Cellulose Composites Reveal Differences in Cellulose Organization after Dilute Acid Pretreatment. *Biomacromolecules* **2019,** *20* (2), 893-903.

135. Vian, B.; Roland, J.-C.; Reis, D.; Mosiniak, M., Distribution and possible morphogenetic role of the xylans within the secondary vessel wall of linden wood. *IAWA Journal* **1992,** *13* (3), 269-282.

136. Mora, F.; Ruel, K.; Comtat, J.; Joseleau, J.-P., Aspect of Native and Redeposited Xylans at the Surface of Cellulose Microfibrils. In *Holzforschung - International Journal of the Biology, Chemistry, Physics and Technology of Wood*, 1986; Vol. 40, p 85.

137. Kabel, M. A.; van den Borne, H.; Vincken, J.-P.; Voragen, A. G. J.; Schols, H. A., Structural differences of xylans affect their interaction with cellulose. *Carbohydrate Polymers* **2007,** *69* (1), 94-105.

138. Köhnke, T.; Östlund, Å.; Brelid, H., Adsorption of Arabinoxylan on Cellulosic Surfaces: Influence of Degree of Substitution and Substitution Pattern on Adsorption Characteristics. *Biomacromolecules* **2011,** *12* (7), 2633-2641.

139. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Pöppler, A.-C.; Brown, S. P.; deAzevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nature Communications* **2016,** *7* (1), 13902.

140. Busse-Wicher, M.; Gomes, T. C. F.; Tryfona, T.; Nikolovski, N.; Stott, K.; Grantham, N. J.; Bolam, D. N.; Skaf, M. S.; Dupree, P., The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of Arabidopsis thaliana. *The Plant Journal* **2014,** *79* (3), 492-506.

141. Pereira, C. S.; Silveira, R. L.; Dupree, P.; Skaf, M. S., Effects of Xylan Side-Chain Substitutions on Xylan–Cellulose Interactions and Implications for Thermal Pretreatment of Cellulosic Biomass. *Biomacromolecules* **2017,** *18* (4), 1311-1321.

142. Martínez-Abad, A.; Berglund, J.; Toriz, G.; Gatenholm, P.; Henriksson, G.; Lindström, M.; Wohlert, J.; Vilaplana, F., Regular Motifs in Xylan Modulate Molecular Flexibility and Interactions with Cellulose Surfaces. *Plant Physiology* **2017,** *175* (4), 1579.

143. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nature Communications* **2019,** *10* (1), 347.

144. Zhu, J. Y.; Wang, G. S.; Pan, X. J.; Gleisner, R., Specific surface to evaluate the efficiencies of milling and pretreatment of wood for enzymatic saccharification. *Chemical Engineering Science* **2009,** *64* (3), 474-485.

145. Yeh, A.-I.; Huang, Y.-C.; Chen, S. H., Effect of particle size on the rate of enzymatic hydrolysis of cellulose. *Carbohydrate Polymers* **2010,** *79* (1), 192-199.

146. Meng, X.; Foston, M.; Leisen, J.; DeMartini, J.; Wyman, C. E.; Ragauskas, A. J., Determination of porosity of lignocellulosic biomass before and after pretreatment by using Simons' stain and NMR techniques. *Bioresource Technology* **2013,** *144*, 467-476.

147. Sun, Q.; Foston, M.; Meng, X.; Sawada, D.; Pingali, S. V.; O'Neill, H. M.; Li, H.; Wyman, C. E.; Langan, P.; Ragauskas, A. J.; Kumar, R., Effect of lignin content on changes occurring in poplar cellulose ultrastructure during dilute acid pretreatment. *Biotechnology for Biofuels* **2014,** *7* (1), 150.

148. Meng, X.; Pu, Y.; Yoo, C. G.; Li, M.; Bali, G.; Park, D.-Y.; Gjersing, E.; Davis, M. F.; Muchero, W.; Tuskan, G. A.; Tschaplinski, T. J.; Ragauskas, A. J., An In-Depth Understanding of Biomass Recalcitrance Using Natural Poplar Variants as the Feedstock. *ChemSusChem* **2017,** *10* (1), 139-150.

149. Yang, B.; Dai, Z.; Ding, S.-Y.; Wyman, C. E., Enzymatic hydrolysis of cellulosic biomass. *Biofuels* **2011,** *2* (4), 421-449.

150. Gao, D.; Chundawat, S. P. S.; Sethi, A.; Balan, V.; Gnanakaran, S.; Dale, B. E., Increased enzyme binding to substrate is not necessary for more efficient cellulose hydrolysis. *Proceedings of the National Academy of Sciences* **2013,** *110* (27), 10922-10927.

151. Liu, Y.-S.; Baker, J. O.; Zeng, Y.; Himmel, M. E.; Haas, T.; Ding, S.-Y., Cellobiohydrolase hydrolyzes crystalline cellulose on hydrophobic faces. *Journal of Biological Chemistry* **2011,** *286* (13), 11195-11201.

152. Chum, H. L.; Johnson, D. K.; Black, S.; Baker, J.; Grohmann, K.; Sarkanen, K. V.; Wallace, K.; Schroeder, H. A., Organosolv pretreatment for enzymatic hydrolysis of poplars: I. Enzyme hydrolysis of cellulosic residues. *Biotechnol Bioeng* **1988,** *31* (7), 643-9.

153. Oinonen, P.; Zhang, L. M.; Lawoko, M.; Henriksson, G., On the formation of lignin polysaccharide networks in Norway spruce. *Phytochemistry* **2015,** *111*, 177-184.

154. Barakat, A.; Winter H Fau - Rondeau-Mouro, C.; Rondeau-Mouro C Fau - Saake, B.; Saake B Fau - Chabbert, B.; Chabbert B Fau - Cathala, B.; Cathala, B., Studies of xylan interactions and cross-linking to synthetic lignins formed by bulk and end-wise polymerization: a model study of lignin carbohydrate complex formation. (0032-0935 (Print)).

155. Mortimer, J. C.; Miles, G. P.; Brown, D. M.; Zhang, Z.; Segura, M. P.; Weimar, T.; Yu, X.; Seffen, K. A.; Stephens, E.; Turner, S. R.; Dupree, P., Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. *Proceedings of the National Academy of Sciences* **2010,** *107* (40), 17409.

156. Terrett, O. M.; Dupree, P., Covalent interactions between lignin and hemicelluloses in plant secondary cell walls. *Current Opinion in Biotechnology* **2019,** *56*, 97-104.

157. Johnson, A. M.; Kim, H.; Ralph, J.; Mansfield, S. D., Natural acetylation impacts carbohydrate recovery during deconstruction of Populus trichocarpa wood. *Biotechnology for Biofuels* **2017,** *10* (1), 48.

158. Fernandes, A. C.; Fontes, C. M.; Gilbert, H. J.; Hazlewood, G. P.; Fernandes, T. H.; Ferreira, L. M., Homologous xylanases from Clostridium thermocellum: evidence for bi-functional activity, synergism between xylanase catalytic modules and the presence of xylan-binding domains in enzyme complexes. *The Biochemical journal* **1999,** *342 (Pt 1)*, 105-10.

159. Donohoe, B. S.; Decker, S. R.; Tucker, M. P.; Himmel, M. E.; Vinzant, T. B., Visualizing lignin coalescence and migration through maize cell walls following thermochemical pretreatment. *Biotechnology and Bioengineering* **2008,** *101* (5), 913-925.

160. Rahikainen, J.; Mikander, S.; Marjamaa, K.; Tamminen, T.; Lappas, A.; Viikari, L.; Kruus, K., Inhibition of enzymatic hydrolysis by residual lignins from softwood—study of enzyme binding and inactivation on lignin-rich surface. *Biotechnology and Bioengineering* **2011,** *108* (12), 2823-2834.

161. Brethauer, S.; Studer, M. H., Biochemical Conversion Processes of Lignocellulosic Biomass to Fuels and Chemicals - A Review. (0009-4293 (Print)).

162. Studer, M. H.; DeMartini, J. D.; Davis, M. F.; Sykes, R. W.; Davison, B.; Keller, M.; Tuskan, G. A.; Wyman, C. E., Lignin content in natural Populus variants affects sugar release. *Proceedings of the National Academy of Sciences* **2011,** *108* (15), 6300-6305.

163. Bhagia, S.; Muchero, W.; Kumar, R.; Tuskan, G. A.; Wyman, C. E., Natural genetic variability reduces recalcitrance in poplar. *Biotechnology for Biofuels* **2016,** *9* (1), 106.

164. Sun, Q.; Khunsupat, R.; Akato, K.; Tao, J.; Labbé, N.; Gallego, N. C.; Bozell, J. J.; Rials, T. G.; Tuskan, G. A.; Tschaplinski, T. J.; Naskar, A. K.; Pu, Y.; Ragauskas, A. J., A study of poplar organosolv lignin after melt rheology treatment as carbon fiber precursors. *Green Chemistry* **2016,** *18* (18), 5015-5024.

165. Pu, Y.; Hu, F.; Huang, F.; Davison, B. H.; Ragauskas, A. J., Assessing the molecular structure basis for biomass recalcitrance during dilute acid and hydrothermal pretreatments. *Biotechnology for Biofuels* **2013,** *6* (1), 15.

166. Chen, F.; Tobimatsu, Y.; Havkin-Frenkel, D.; Dixon, R. A.; Ralph, J., A polymer of caffeyl alcohol in plant seeds. *Proceedings of the National Academy of Sciences* **2012,** *109* (5), 1772.

167. Morrison, D.; van Dyk, J. S.; Pletschke, B. I., The effect Of alcohols, lignin and phenolic compounds on the enzyme activity of Clostridium Cellulovorans XynA. *BioResources; Vol 6, No 3 (2011)* **2011**.

168. Mayer-Laigle, C.; Blanc, N.; Rajaonarivony, R. K.; Rouau, X., Comminution of Dry Lignocellulosic Biomass, a Review: Part I. From Fundamental Mechanisms to Milling Behaviour. *Bioengineering (Basel)* **2018,** *5* (2), 41.

169. Zhu, J. Y., Physical Pretreatment − Woody Biomass Size Reduction − for Forest Biorefinery. In *Sustainable Production of Fuels, Chemicals, and Fibers from Forest Biomass*, American Chemical Society: 2011; Vol. 1067, pp 89-107.

170. Tassinari, T.; Macy, C.; Spano, L.; Ryu, D. D. Y., Energy requirements and process design considerations in compression-milling pretreatment of cellulosic wastes for enzymatic hydrolysis. *Biotechnology and Bioengineering* **1980,** *22* (8), 1689-1705.

171. McMillan, J. D., Pretreatment of lignocellulosic biomass. ACS Publications: 1994.

172. Alvo, P.; Belkacemi, K., Enzymatic saccharification of milled timothy (Phleum pratense L.) and alfalfa (Medicago sativa L.). *Bioresource Technology* **1997,** *61* (3), 185-198.

173. Behera, S.; Arora, R.; Nandhagopal, N.; Kumar, S., Importance of chemical pretreatment for bioconversion of lignocellulosic biomass. *Renewable and sustainable energy reviews* **2014,** *36*, 91-106.

174. Ackerson, M.; Ziobro, M.; Gaddy, J. L. In *Two-stage acid hydrolysis of biomass*, 1981.

175. Kootstra, A. M. J.; Beeftink, H. H.; Scott, E. L.; Sanders, J. P. M., Comparison of dilute mineral and organic acid pretreatment for enzymatic hydrolysis of wheat straw. *Biochemical Engineering Journal* **2009,** *46* (2), 126-131.

176. Jung, Y. H.; H., K. K., Chapter 3. Acidic Pretreatment. In *Pretreatment of Biomass*, Pandey, A.; Negi, S.; Binod, P.; Larroche, C., Eds. Elsevier: Amsterdam, 2015; pp 259-264.

177. Jönsson, L. J.; Alriksson, B.; Nilvebrant, N.-O., Bioconversion of lignocellulose: inhibitors and detoxification. *Biotechnology for Biofuels* **2013,** *6* (1), 16.

178. Hu, F.; Ragauskas, A., Pretreatment and lignocellulosic chemistry. *Bioenergy Research* **2012,** *5* (4), 1043-1066.

179. Agbor, V. B.; Cicek, N.; Sparling, R.; Berlin, A.; Levin, D. B., Biomass pretreatment: Fundamentals toward application. *Biotechnology Advances* **2011,** *29* (6), 675-685.

180. Esteghlalian, A.; Hashimoto, A. G.; Fenske, J. J.; Penner, M. H., Modeling and optimization of the dilute-sulfuric-acid pretreatment of corn stover, poplar and switchgrass. *Bioresource Technology* **1997,** *59* (2), 129-136.

181. Kim, Y.; Kreke, T.; Mosier, N. S.; Ladisch, M. R., Severity factor coefficients for subcritical liquid hot water pretreatment of hardwood chips. *Biotechnology and Bioengineering* **2014,** *111* (2), 254-263.

182. Selig, M. J.; Viamajala, S.; Decker, S. R.; Tucker, M. P.; Himmel, M. E.; Vinzant, T. B., Deposition of Lignin Droplets Produced During Dilute Acid Pretreatment of Maize Stems Retards Enzymatic Hydrolysis of Cellulose. *Biotechnology Progress* **2007,** *23* (6), 1333-1339.

183. Samuel, R.; Pu, Y.; Foston, M.; Ragauskas, A. J., Solid-state NMR characterization of switchgrass cellulose after dilute acid pretreatment. *Biofuels* **2010,** *1* (1), 85-90.

184. Kumar, R.; Wyman, C. E., Access of cellulase to cellulose and lignin for poplar solids produced by leading pretreatment technologies. (1520-6033 (Electronic)).

185. Foston, M.; Ragauskas, A. J., Changes in lignocellulosic supramolecular and ultrastructure during dilute acid pretreatment of Populus and switchgrass. *Biomass and Bioenergy* **2010,** *34* (12), 1885-1895.

186. Li, C.; Knierim, B.; Manisseri, C.; Arora, R.; Scheller, H. V.; Auer, M.; Vogel, K. P.; Simmons, B. A.; Singh, S., Comparison of dilute acid and ionic liquid pretreatment of switchgrass: Biomass recalcitrance, delignification and enzymatic saccharification. *Bioresource Technology* **2010,** *101* (13), 4900-4906.

187. Bali, G.; Meng X Fau - Deneff, J. I.; Deneff Ji Fau - Sun, Q.; Sun Q Fau - Ragauskas, A. J.; Ragauskas, A. J., The effect of alkaline pretreatment methods on cellulose structure and accessibility. (1864-564X (Electronic)).

188. Kim, S.; Holtzapple, M. T., Lime pretreatment and enzymatic hydrolysis of corn stover. *Bioresource Technology* **2005,** *96* (18), 1994-2006.

189. Gould, J. M., Alkaline peroxide delignification of agricultural residues to enhance enzymatic saccharification. (0006-3592 (Print)).

190. Pandey, A.; Negi, S.; Binod, P.; Larroche, C., *Pretreatment of Biomass: Processes and Technologies*. Elsevier Science: 2018.

191. Soares Rodrigues, C. I.; Jackson, J. J.; Montross, M. D., A molar basis comparison of calcium hydroxide, sodium hydroxide, and potassium hydroxide on the pretreatment of switchgrass and miscanthus under high solids conditions. *Industrial Crops and Products* **2016,** *92*, 165-173.

192. Yoon, H. H.; Wu, Z. W.; Lee, Y. Y., Ammonia-recycled percolation process for pretreatment of biomass feedstock. *Applied Biochemistry and Biotechnology* **1995,** *51* (1), 5-19.

193. Kim, T. H.; Lee, Y. Y., Pretreatment of corn stover by soaking in aqueous ammonia. (0273-2289 (Print)).

194. Balan, V.; Bals B Fau - Chundawat, S. P. S.; Chundawat Sp Fau - Marshall, D.; Marshall D Fau - Dale, B. E.; Dale, B. E., Lignocellulosic biomass pretreatment using AFEX. (1940-6029 (Electronic)).

195. Salapa, I.; Katsimpouras, C.; Topakas, E.; Sidiras, D., Organosolv pretreatment of wheat straw for efficient ethanol production using various solvents. *Biomass and Bioenergy* **2017,** *100*, 10-16.

196. Zhao, X.; Cheng, K.; Liu, D., Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis. *Applied Microbiology and Biotechnology* **2009,** *82* (5), 815.

197. Zhu, S.; Wu, Y.; Chen, Q.; Yu, Z.; Wang, C.; Jin, S.; Ding, Y.; Wu, G., Dissolution of cellulose with ionic liquids and its application: a mini-review. *Green Chemistry* **2006,** *8* (4), 325-327.

198. Zhao, H.; Jones, C. L.; Baker, G. A.; Xia, S.; Olubajo, O.; Person, V. N., Regenerating cellulose from ionic liquids for an accelerated enzymatic hydrolysis. *Journal of Biotechnology* **2009,** *139* (1), 47-54.

199. Sindhu, R.; Binod, P.; Pandey, A., Biological pretreatment of lignocellulosic biomass – An overview. *Bioresource Technology* **2016,** *199*, 76-82.

200. Janusz, G.; Pawlik, A.; Sulej, J.; Swiderska-Burek, U.; Jarosz-Wilkolazka, A.; Paszczynski, A., Lignin degradation: microorganisms, enzymes involved, genomes analysis and evolution. *FEMS Microbiol Rev* **2017,** *41* (6), 941-962.

Chapter 3). Characterization of lignocellulosic biomass:

In Chapter 3, the chemical (composition, linkage molecular weight) and physical (a form of cellulose) structure of individual cell wall polymers such as cellulose, hemicellulose, pectin, and lignin were discussed. Additionally, conceptual models of plant cell wall based on the type of polymer interactions and distribution of these polymers in the cell wall were shown. As lignocellulosic biomass is a multicomponent, heterogeneous and complex system, the information to build the model has been obtained by using various techniques and probes. Over the past four decades, tremendous research effort has been made to develop different techniques that cover length scales spanning from several hundred nanometers to as small as angstroms[1] (Figure 3.1).

Scattering techniques, using X-rays or neutrons, are often used as complementary tools to chemical characterization and microscopy to gain insights into the structural features of lignocellulosic biomass. Information such as the overall cell wall morphology, distribution of lignin, conformation of individual cell wall polymers in solution or dried state, and the size of cellulose microfibrillar cross-section, its hierarchical structure and structure of the crystallites within the cellulose can be obtained with scattering. Minimum sample manipulation needs to be done to the biomass sample which makes scattering one of the few techniques through which biomass can be studied without any modifications.[1] Additionally, variety of sample environments are available due to which the biomass structure can be studied at different temperature, pressure, humidity conditions in real-time.[2] We combined small-angle scattering (SAS) and wide-angle scattering (WAS) data of the plant cell wall to gain information in size ranges of several hundred nanometers to a few angstroms.

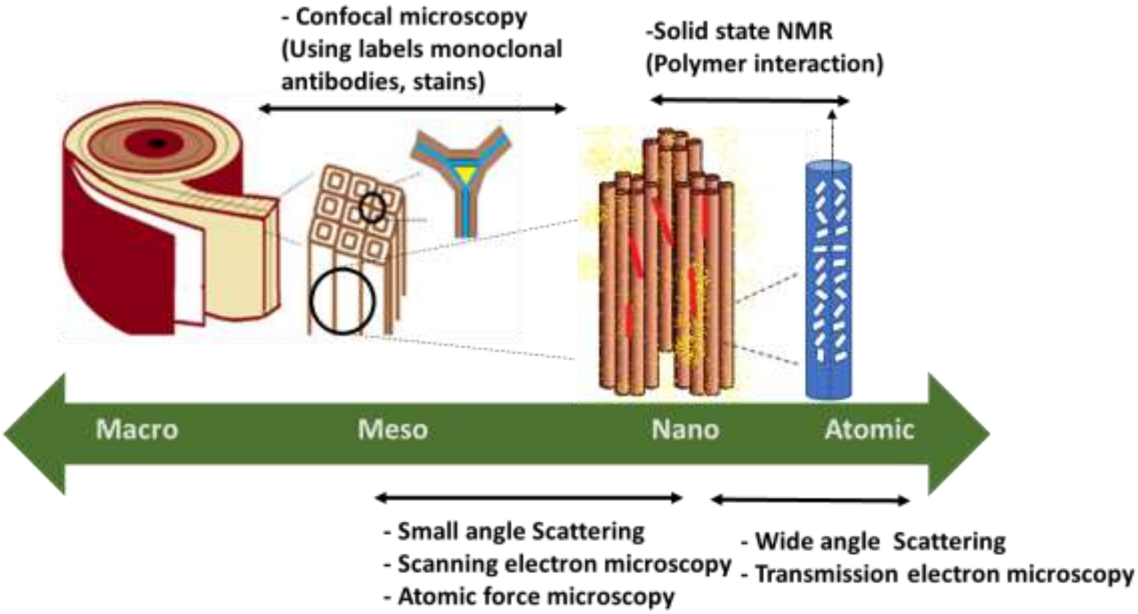


Figure 3.1: Compilation of the few of the techniques and probes that have been used to gain deeper insights into the arrangement and interaction of the polymers in the plant cell wall

3.1). Fundamental of scattering[3-4]

When an incident beam interacts with matter, most of them are transmitted through the matter. However, there is a fraction of the incident beam that is absorbed and transformed into other forms of energy and a small fraction that gets scattered in directions of propagation other than the incident beam. The scattering event in which the energy of incident radiation is conserved, and there is a change in the direction of propagation of the scattered radiation is called Thompson scattering or elastic scattering. The change is momentum is described by a momentum transfer vector, also called scattering vector Q (also referred in the literature as vector q or vector s) (Figure 3.2).

The scattering vector, Q can be calculated by knowing the angle between the transmitted beam and scattered beam (scattering angle, Θ) and the wavelength of the applied radiation (Eq.3.1).

$$Q = \frac{4\pi \times \sin(\Theta)}{\lambda} \qquad \text{Eq. 3-1}$$

The unit of Q is inverse length and due to this, the scattering image shows "form of the object" in reciprocal space. The Q-range defines the span of length covered in a typical scattering experiment. For instance, small-angle scattering (typically used for lignocellulosic biomass) covers length scale $0.001\ \text{Å}^{-1} < Q < 0.6\ \text{Å}^{-1}$ which correspond to a real space dimension of approximately 600 nm down to 1 nm while wide-angle scattering (for cellulose) $0.5\ \text{Å}^{-1} < Q < 3\ \text{Å}^{-1}$.

The scattering events in which energy loss occurs due to which the scattered beam has less energy than the incident beam is called inelastic scattering or Compton scattering. Here, the scattered radiation has a different wavelength as compared to the incident radiation. It cannot produce an interference pattern and carries no information about the structure of the matter. In the case of small and wide-angle scattering, it contributes to increasing the background noise.

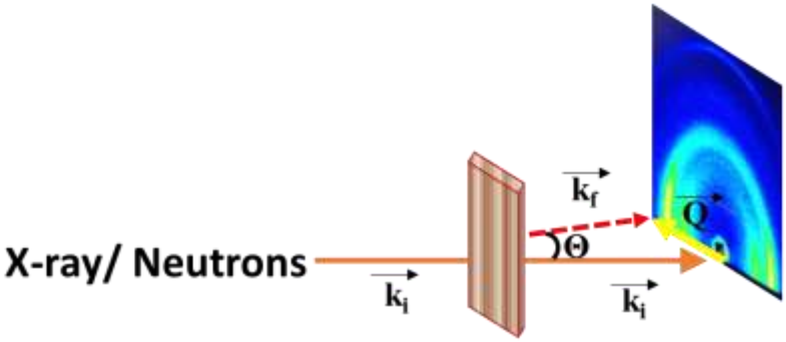


Figure 3.2: Incident radiation (X-ray or neutron) hitting a wood chip in transmission mode. The intensity of the scattered radiation is detected on an area detector. The difference in momentum of the incident radiation (k_i) and scattered radiation (k_f) is measured by vector, Q

In matter, many particles interact with the incoming radiation and act as centers for scattering the radiation. For instance, if the incoming radiation is X-rays then atoms in the matter would act as centers, and if neutrons are the source of radiation then nuclei of the atom in the matter will scatter incoming neutrons. The scattered X-ray or neutrons from each scattering center travel as spherical waves. In case of elastic scattering, the waves from neighboring scattering center would cause an interference pattern as shown in Figure 3.3.

The waves reach the detector and depending on how constructive the interference was, a spot with varying degrees of intensity can be seen on the detector. Each matter has a particular arrangement of the scattering center due to which the interference pattern gives characteristic information about the structure of the objects in that matter. These scattering centers can be densely packed like cellulose microfibrils in the plant cell wall or glucan chain in the crystallites of cellulose microfibrils or they can be present in dilute solution like extracted hemicellulose or lignin dissolved in solvent. If the centers are densely packed, then they give rise to structure factors in the scattering pattern, while from most dilute solutions the shape and size of the particles can be determined.

The interference of the waves can be constructive if the phase of the waves is synchronous, destructive if the difference between the phase of waves is exactly 90° and somewhere in between depending on the difference in the phase of each wave. Each scattered wave also has an amplitude and a wave propagating in a given direction at a certain time t can mathematically be described as follows (Eq. 3.2):

$$E = A\sin(2\pi\vartheta t - \emptyset) \qquad \text{Eq. 3-2}$$

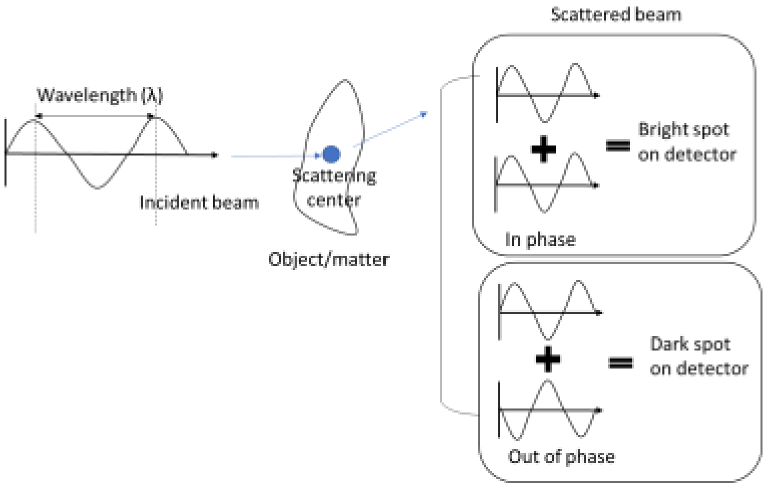


Figure 3.3: Incident beam gets scattered due to interaction with particles in the matter (scattering centers) and if the scattered waves from neighboring particles are in phase, they will produce a bright spot on the detector. If the waves are completely out of phase, they will form dark spot on the detector. Note if the in-phase waves are produced from close neighbors the intensity on the detector will be brighter then when they are from distant particles[5]

where, E for the incoming electromagnetic radiation is the electric field intensity, ϑ is the frequency of the wave and $\emptyset$ is the phase of the wave

It is possible to have scattering centers/particles of different kinds in which case the amplitude and phase of each wave will be different. In such a case, complex number (a+ib) are used to represent the wave in a complex plane and the analytical expression of a particular wave is now a complex number (A cosΦ +i A sin Φ) and this is equal to e^{ix} using the power-series expansion.

Since the scattered wave is spherical, the wave function (ψ_s) can be represented by the Eq. 3.3:

$$\psi_s = -\frac{b}{r}e^{-ikr} \qquad \text{Eq. 3-3}$$

Where, b is the scattering length of the scattering center and is the measure of the scattering ability of the scattering center after interaction with the incident, k is the momentum transfer vector and r is the distance between the scattering center and detector.

The interference pattern from a macroscopic sample are formed from all the scattering center in the macroscopic sample. It is important to note that the fundamental quantity determined in any scattering experiment is the differential scattering cross-section ($d\sigma_s/d\Omega$), which is defined as the probability of scattering event that will occur in the elemental solid angle dΩ

Considering the scattering from N number of scattering centers having scattering length b within the sample, the differential scattering cross-section, $d\sigma/d\Omega$ can be expressed in terms of scattering vector Q as shown in Eq. 3.4 below, which provides the structural information

$$\frac{d\sigma}{d\Omega}(Q) = \frac{1}{N}\left|\sum_i^N b_i e^{i\vec{Q}.\vec{r}}\right|^2 \qquad \text{Eq. 3-4}$$

The distribution of particles in a given volume is inhomogeneous and that is accounted for by the Eq. 3.5 where ρ(r) parameter is the summation of all the scattering length (b_i) in a given volume ($\bar{V}$) in the distance between the scattering center in the volume to the detector (r) and is called the scattering length density distribution

$$\rho(\mathrm{r}) = \frac{\Sigma_i^n b_i}{\bar{V}}(\mathrm{r}) \qquad \text{Eq. 3-5}$$

Substituting Eq. 3.5 in Eq. 3.4 and integrating over the scattering length density distribution across the entire sample after normalizing by volume gives Eq. 3.6:

$$\frac{d\Sigma}{d\Omega}(Q) = \frac{N}{V}\frac{d\sigma}{d\Omega}(Q) = \frac{1}{V}\left|\int_V \rho\,(r)e^{iQ.r}dr)\right|^2 \qquad \text{Eq. 3-6}$$

The intensity I(Q) is a shorthand notation for the macroscopic scattering cross-section dΣ/dΩ. Hence, the Eq. 3.7 can also be written as follows:

$$I\,(Q) = \frac{d\Sigma}{d\Omega}(Q) = \frac{N}{V}\frac{d\sigma}{d\Omega}(Q) = \frac{1}{V}\left|\int_V \rho\,(r)e^{iQ.r}dr)\right|^2 \qquad \text{Eq. 3-7}$$

<u>3.1.1). Contrast variation</u>

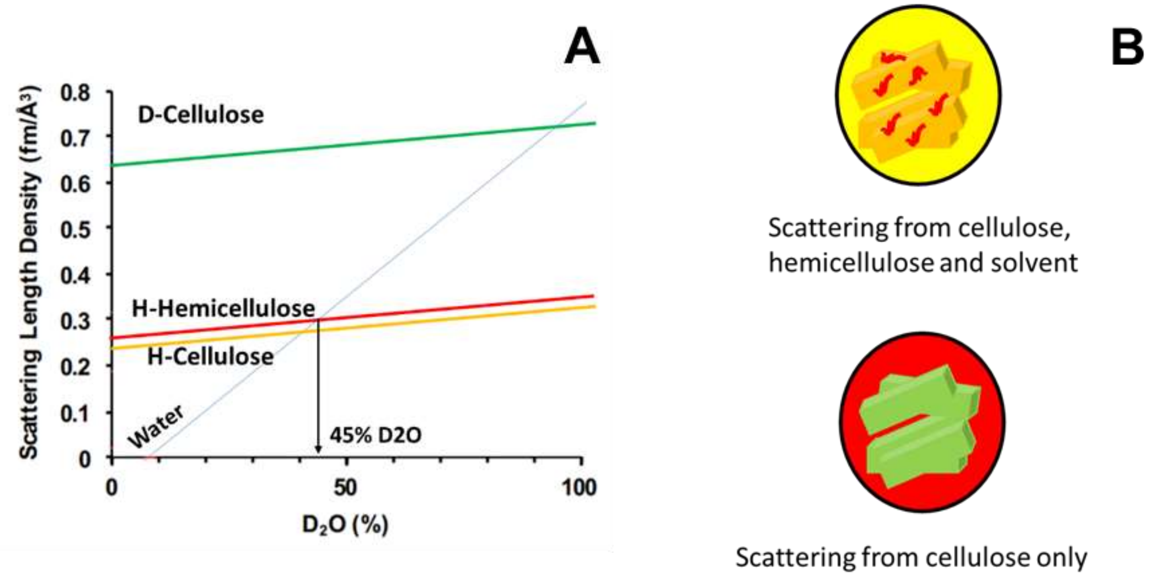
The quantity ρ(r) is of particular interest for the multi-component system. In order to study the multi-component system such as lignocellulosic biomass, the ability to vary the scattering length density through contrast variation (such as the hydrogen-deuterium exchange) is a key to study the various structural features. For instance, if we consider the cellulose microfibrils as one scattering phase with scattering length density ρ_1 and hemicellulose with water as other scattering phase with density ρ_2 then Eq. 3.8 becomes as follows:

$$I\,(Q) = (\rho_1 - \rho_2)^2\,\frac{1}{V}\left|\int_V e^{iQ.r}dr)\right|^2 \qquad \text{Eq. 3-8}$$

The scattering length density is dependent on the scattering length *b* which in turn depends on the type of atoms and the type of incident radiation. X-rays interact with the electron cloud of an atom while the neutrons interact with nuclei due to this a difference in *b*. Most biopolymers are organic in nature and majority of them have an elemental composition of carbon, hydrogen, and oxygen. Due to this, if the scattering of a deuterated biopolymer is measured in a hydrogenated media with neutrons a difference in scattering length density also called "contrast". The deuterated polymer can be put into different ratios of H:D containing solvent and the contrast between the different phases can be varied. As seen in Figure 3.4, hydrogenated cellulose and hydrogenated hemicellulose have very similar scattering length density due to which there is no contrast and the measured scattering curve will have contributions from both the polymers. If both polymers are placed in 45% D_2O, the scattering curve from 45% D_2O is same as that from H-cellulose and H-hemicellulose and this is called "contrast match." However, if D-cellulose is put in 45% D_2O then due to the difference in the scattering length density, scattering contributions from D-cellulose can be highlighted.

3.1.1). Form factor[6]

This intensity measured at the detector is the square of the amplitude of the scattered wave. Summing up all the wave amplitudes results in a SAS pattern. In the SAS pattern, the phase (Φ) information needed to point out the exact location of an atom is lost. This makes small-angle scattering different from electron microscopy where the phase information can be used and an image of the object can be obtained. The SAS pattern does contain information about the particle shape and size which is also called the form factor (P(Q)). A Fourier transform function is applied to particle shape (such as sphere, cylinder, etc.) and size in real space to obtain a modeled scattering pattern in reciprocal space.



Scattering from cellulose, hemicellulose and solvent

Scattering from cellulose only

Figure 3.4: A). A plot showing the scattering length density (if incident beam is neutron radiation) of different ratio of H_2O and D_2O (blue line), hydrogenated (H) cellulose in different ratio of H_2O: D_2O solvent (orange line), H- hemicellulose in different ratio of H_2O: D_2O solvent (redline),and deuterated (D) cellulose in different ratio of H_2O: D_2O solvent (green line) . The downward arrow indicates the scattering length density of 45% H_2O containing solvent and hydrogenated cellulose or hemicellulose is similar. B). (top) Schematic showing not much difference in contrast between the cellulose (orange), hemicellulose (red) and 45 % H_2O (yellow) while (bottom) D-cellulose in 45% D_2O has a different contrast. Note that hemicellulose is not seen as it has the same contrast as 45% D_2O due to which the scattering intensity is similar to 45% D_2O

The modeled scattering patterns are scaled to fit the experimental data and a match is obtained. In order to make a reasonable assumption to fit a particular model to the experimental scattering pattern prior knowledge of the particle with other techniques such as microscopy is very helpful.

3.1.2). Structure factor[7]

When particles are densely packed and are arranged in a particular order, the scattering pattern will have contributions not just from the particle but also the distance between adjacent particles. In this case, besides the form factor there is additional interference from between the particles that contributes to the observed intensity. This additional interference is called the structure factor which is multiplied to the form factor and together they form the SAS pattern. When the particles are highly ordered and have a periodic arrangement (for instance in the crystalline region of cellulose) a peak is seen in the scattering pattern. This peak represents the distance between the particles or the distance between the plane that have the periodically arranged particle (d-spacing) which can be calculated using Bragg's Law (Eq. 3.9).

$$n\lambda = 2d\sin\Theta \qquad \text{Eq. 3-9}$$

Where, λ is the wavelength of the incident radiation, d is the distance between the planes and Θ is the scattering angle.

The same equation can be expressed in terms of Q by substituting Eq. 3.10 into the above-mentioned equation

$$d_{Bragg} = \frac{2\pi}{Q_{peak}} \qquad \text{Eq. 3-10}$$

In summary, the scattering intensity I(Q) represents the interference of scattered photons/neutrons either from different positions of the same object (form factor P(Q)) and/or the interference of photons/neutrons scattered from different objects which are periodically arranged (structure factor, S(Q)). So, the measured intensity in Eq. 3.11 (corrected for background and put on an absolute scale, refer section 4.2) can also be expressed as

$$\mathrm{I(Q)} = \frac{\mathrm{N}}{\mathrm{V}} \rho^2 \mathrm{V_p}^2 \ \mathrm{P(Q)}S(\mathrm{Q}) \qquad \text{Eq. 3-11}$$

Where, N is the number density of the particles, V is the sample irradiation volume, Vp is the volume of the scattering particle, ρ is the scattering length density, P(Q) is the form factor and S(Q) is the structure factor

3.2). Scattering data collection and processing

3.2.1). Scattering instrument[3]

Figure 3.5 shows the schematic showing the basic components of a small angle scattering instrument. The source produces beam (X-ray or neutrons) with various wavelengths which pass through a monochromator. The monochromator only allows a particular wavelength of beam to pass through. The beam with a certain wavelength then passes through a collimation system. The pinhole or slit geometry of the collimators makes the beam narrow and define the beam shape. The collimated beam then irradiates the sample which is placed in the sample holder. The scattered beam then gets collected at the detector. The transmitted beam that goes through the sample is intense and can damage the detector. To avoid damage to the detector a beam stop is positioned to stop the transmitted beam. At times the beam stop is semi-transparent and this allows the collection of transmitted and scattered data in one shot. The position of the detector with respect to the sample determines the measured Q-range.

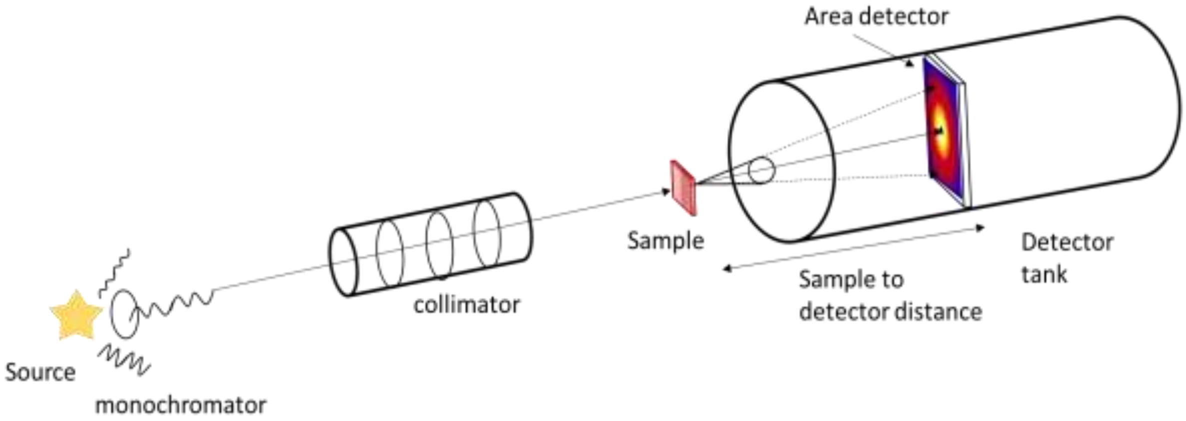


Figure 3.5: Assembly of the different components of a scattering instrument. The source of radiations can be x-rays or cold neutron. Desired wavelength is selected with a monochromator or a velocity selector

3.2.2). Sample environment

The samples for the neutron scattering experiments were packed in quartz cuvettes or enhanced angle pressure cell (Figure 3.6 panel A). The thickness of both the cells was 1mm.

A few things should be considered before measuring the samples. These include finding the optimum Q-Range, calibrating the sample to detector distance, determining the detector efficiency and measuring background intensity (empty cell scattering, blocked beam)

3.2.3). Data processing[5]

The scattered waves are recorded on a two-dimensional (2D) area detector and the resulting intensity pattern would be 2D scattering intensity of the particle. It is represented as I(q_x, qy) where intensity is along the x and y pixel direction. If the particles of the scattered sample have no order or orientation it would result in an isotropic pattern. If there is a particle defined order in the sample then it would form an anisotropic 2D intensity pattern. In case of isotropic scattering, the 2D data is converted into 1D scattering by radially averaging in the azimuthal direction for each Q-value. The 2D data can also be box-averaged or pie averaged. The resolution in Q is decided by binning the number of pixels. Using the sample to detector distance and wavelength, the pixels are converted to Q and a plot of I(Q) vs Q, also called 1D scattering profile.

Along with the scattering from the sample, the medium in which the sample is placed also contributes to the scattering. while the empty cell scattering can be subtracted, there is at times the media such as a solvent in which the sample is present. The sample and solvent scattering may not be the same. If the sample absorbs more than the solvent then the resulting intensity of the sample will be negative values rendering the experiment meaningless.

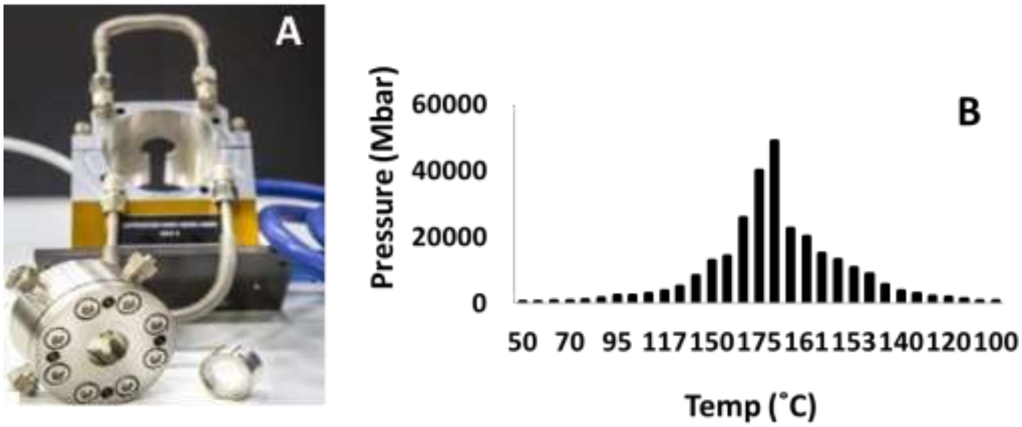


Figure 3.6: A). The assembled enhanced angle pressure cell with its holder.[2] The pressure cell has two sapphire windows and a spacer (seen next to the cell) that modulated the thickness of sample. B). Pressure profile of water filled in the pressure cell during heating phase from 25 ˚C to 180 ˚C and cooling phase of 180 ˚C to 25 ˚C

So, the experimental scattering curve of both the sample and solvent must be scaled according to their transmission and the difference of that would be intensity from the scattering particle of the sample. Intensities are reported on an absolute scale which is achieved by measuring two extra samples. One of them is either a standard whose scattering intensity is known such as water and the other is the empty cell. The background corrected sample intensity is divided by the mean intensity of the reference material. This results in the sample intensity being on absolute scale which is suitable for molecular weight determination and comparison of scattering data from different experiments.

3.3). Data analysis

The data analysis of the scattering curves is complex and the choice of an appropriate model is critical to allow and accurate quantification and interpretation of the observed structural features. Here I discuss the several common approaches I used for analyzing small and wide-angle scattering curves.

3.3.1). Guinier analysis[6]

For some systems, it is not possible to know the exact shape of the particle. In these cases, the Guinier approximation is applied in which any form factor P(Q) at small angles can be considered to have a Gaussian curve. The Gaussian curvature is due to the overall size of the particle and so using this the radius of gyration (Rg) of the particle can be determined. The Guinier approximation is as follows,

$$\mathrm{I(Q)} = \mathrm{I_0} \times \mathrm{e}^{-\frac{\mathrm{Q^2 \times Rg^2}}{3}} \qquad \text{Eq. 3-12}$$

This equation can also be expressed as follows,

$$\ln(\mathrm{Q}) = \ln(\mathrm{I}_0) - \frac{\mathrm{Rg}^2\mathrm{Q}^2}{3} \qquad \text{Eq. 3-13}$$

The above equation is like that of a straight line and so on a plot of ln (Q) vs Q^2 the slope would be $-\frac{Rg^2}{3}$ Such a plot is called the Guinier plot and is widely used to determine the radius of gyration, Rg. The Rg can also be determined for values of Q where a condition $Q \times R_g \ll 1$ is met. The Rg is especially useful for hierarchical systems like bacterial cellulose where the shape of macrofibril is not accurately known. However, if the shape of the particle is known and the particle is known to have a uniform density then the following equations can be applied to know the size of the particle.

For sphere with radius R, $R_g = \sqrt{\frac{3}{5}}R$

For cylinder with length L and cross-sectional radius R, $R_g = \left(\sqrt{\frac{L^2}{12} + \frac{R^2}{2}}\right)$

3.3.2). Porod's law[5]

Besides determining the size of a particle, the surface (rough or smooth) of the particle can also be determined from a SAS 1 D curve. Additionally, a system in which there are interacting particles a cluster formation is observed. The particles depending on the degree of crosslinking may form densely packed or loose clusters. The so formed network structure can be determined and is called the mass fractal (Figure 3.7).

According to Porod's law, the scattering profile is plotted as ln I(Q) vs Q would have a particle size feature towards low Q followed by a linear decay towards highQ and would have the following equation (Eq. 3.14),

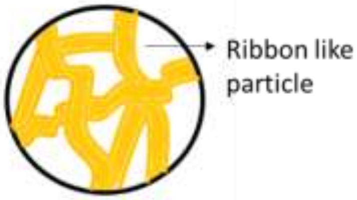


Figure 3.7: Schematic shows ribbons of the bacterial cellulose. Each ribbon is a scattering particle and they cross each other to form a network like structure. The mass fractal would represent the degree of crosslinking or packing (loose or dense) of ribbons to form the network. The interface the ribbon makes with the surrounding would give an indication of the degree of roughness of the surface or surface fractal of the ribbon

$$I(Q) \approx Q^{-\alpha} \qquad \text{Eq. 3-14}$$

Where, α is the power law exponent

The slope obtained by fitting the linear decay can be used to know local structural features of the particle such as surface or mass fractal. In cases where the particles are not forming a network and are monodisperse, the power law exponent indicates the shape of the particle. For instance, n =1 for rigid rod; n=2 for disk

3.3.3). Unified fit for hierarchical structure[8]

Complex systems with many components will have several structural features ranging from micrometer to nanometer and contribute from all the features that will be present in the small-angle scattering curve. Additionally, at times one feature may arise due to contribution of two or more component and hence a model that would help in understanding the relationship between related structural features and their individual contribution at different length scales needs to be used to fit the scattering curve. The scattering curve may have more than one particle size feature along with power law scattering. Especially in lignocellulosic biomass, it is difficult to judge the contribution of one component in isolation, fitting parts of the curve with Guinier -Porod or form factors may not be a correct approach. Beaucage et al. developed a fitting approach in which the scattering curve is divided into individual levels which can be fit with Guinier and Porod law. Each of the levels would have an effect in it's preceding or succeeding level and they are summed to give a fit for the overall curve which is called the Unified fit.

An example of small-angle scattering curve of poplar wood chip fit with the Unified fit is shown in Figure 3.8.

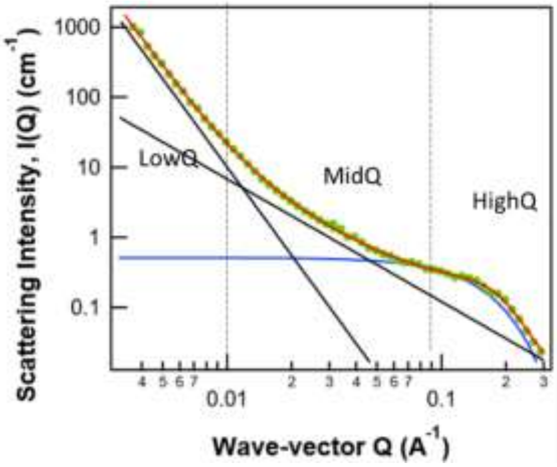


Figure 3.8: Experimentally obtained scattering curve of poplar (green), which is fit using the Unified fitting approach (red). The dashed vertical lines show the levels in which individual Guinier (blue) and porod (black) where fit.

Scattering is mainly from the cell wall of poplar and as the cell wall is a complex material the scattering curve has structural information spanning length scale from about 1μm to 10Å. Scattering analysis that describes hierarchical structures: (I) Q-range 0.001Å^{-1} to 0.01 Å^{-1} shows the overall cell wall morphology (Porod law exponent ~4); (II) Q-range- 0.01Å^{-1} to 0.08 Å^{-1} is due to mass fractal scattering from the matrix cell wall polymer (Porod law exponent ~2.1); (III) Q-range- 0.08Å^{-1} to 0.2 Å^{-1} scattering from cross-section of cellulose microfibrils and distance between the microfibrils; (IV) Q-range- 0.2Å^{-1} to 0.5 Å^{-1} has scattering from the primary nanoparticles with a sharp interface (high Q Porod slope of -4) (V) Wide angle scattering from atomic structure (not shown).

While the Unified fit provides the flexibility to fit the entire Q-range, to come to a conclusion that a feature is due to a certain component needs support from other techniques. However, the unified approach often offers the only reasonable approach to understanding small-angle scattering. It also offers the opportunity to resolve scattering features obscured by the overlap of structural levels

3.3.4). Wide angle scattering data analysis

3.3.4.1). Cellulose crystallinity[9-10]

Crystallinity determination from X-ray diffraction curve of cellulose depends on the way the amorphous phase is determined. Figure 3.9 shows the three methods for estimating the amorphous content and determining percentage crystallinity. These include peak height method in which the height of amorphous peak (2Θ of 18.3˚) and height of crystalline peak (200) are determined and a ratio of these heights is calculated which equates to crystallinity.

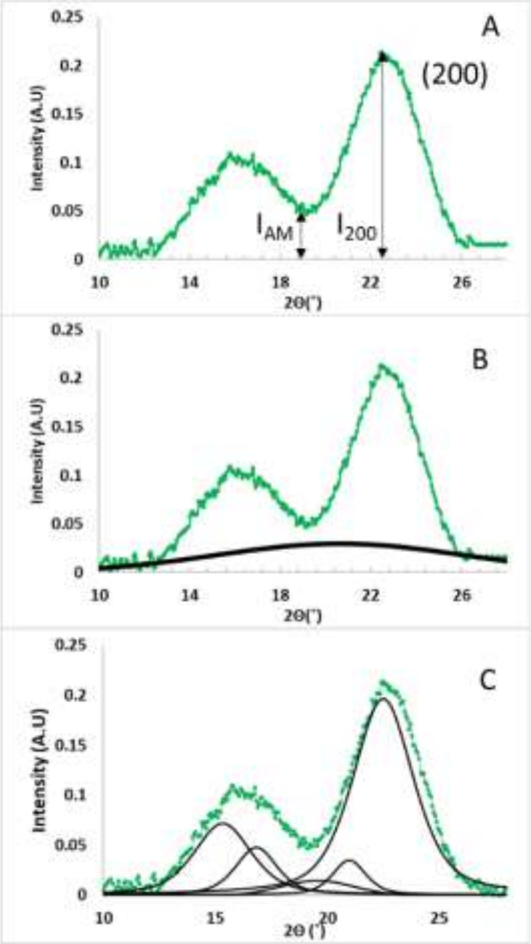


Figure 3.9: The three most common methods for determining crystallinity from X-ray diffraction spectra: Peak height method (A), amorphous subtraction method (B) and peak deconvolution method (C). The diffraction spectrum was obtained by measuring poplar wood chip

Another method for determining crystallinity is by deconvoluting the amorphous and crystalline region done by fitting the curve with pseudoVoight, gaussian and Voight peaks. In this method, the area under the amorphous peak is subtracted from the crystalline peak area of the highest peak (200). The crystallinity is determined from the ratio of the area of all crystalline peaks to the total area. The third method is the background subtraction method in which the amorphous region is determined from the diffraction spectra of amorphous standard. This area is subtracted from the total scattering area to give the crystalline area. The crystallinity is the ratio of crystalline area and the total scattering area. All these three methods can be used to get crystallinity values between a set of samples and obtain a trend. The crystallinity values are not absolute and cannot be compared with a different sample set.

3.3.4.2). Crystallite size[7]

The dimension of the crystallite size can also be determined from the 1D scattering pattern. Using the Scherrer equation the crystallite dimension along each reflection plane can be calculated using the following Eq.3.15

$$\mathrm{t} = (0.9 \times \lambda) \div (\beta \times \cos \theta) \qquad \text{Eq. 3-15}$$

where, t is the size of the crystallite, λ is the wavelength of the incident X-rays, β and 2θ are the values for FWHM and peak position, respectively, obtained as described above.

3.3.4.3). Crystallite orientation[11]

Particles that are not spherical may have a fixed alignment about a certain axis which can be measured with Herman's orientation parameter (f). Cellulose crystallites with the cellulose microfibrils are one such example and Figure 3.10 shows the 2D diffraction pattern of poplar in which the plane (200) has strong equatorial scattering. The crystallite orientation can be calculated from the intensity profile obtained by along the circular path that goes through the equatorial reflection and is centered at the primary beam position. The angle formed from the center to any point on the circular path is called the azimuthal angle (Φ). The following equations (3.16 and 3.17) are used to calculate the Herman's orientation parameter (*f*) after generating the I(Φ) vs Φ profile.

$$\langle \cos^2 \Phi \rangle = \frac{\int_0^{\pi} \mathrm{I}(\emptyset) \cos^2 \emptyset \sin \emptyset \ \mathrm{d}\emptyset}{\int_0^{\pi} \mathrm{I}(\emptyset) \sin \emptyset \ \mathrm{d}\emptyset} \qquad \text{Eq. 3-16}$$

$$\mathrm{f} = \frac{3 \langle \cos^2 \Phi \rangle - 1}{2} \qquad \text{Eq. 3-17}$$

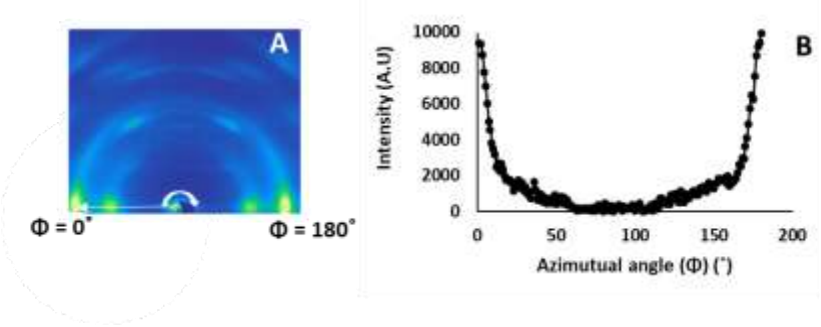


Figure 3.10: A) 2D scattering pattern showing the crystallites have preferred orientation as there are strong equatorial peaks. B). a circular path with certain radius that connects the primary center of the beam to the equatorial peak is drawn and the intensity along the path is plotted against the angle each point on the circular path makes with the beam center (Φ)

3.4). Quantitative saccharification

The wood chips are freeze-dried to constant weight (typically 24 h) and 50 mg of each sample was added to a 15 ml Pyrex heavy wall pressure vessel (Chemglass, CG-1880, USA). The monosaccharide composition was determined based on a previously reported procedure [12-13]. Each sample was incubated in 0.5 ml of 72% (w/w) H_2SO_4 at 30 ˚C for 1 h. H_2O was then added to reach a final H_2SO_4 concentration of 4% (w/w), followed by heating in a commercial autoclave (Panasonic, MLS3781-L) at 121˚C and 2.32 atm for 1 h. The solution was neutralized to pH~7 using calcium carbonate and filtered before loading to AminexTMHPX-87P (Bio-Rad Laboratories Inc., Hercules, CA) column for determination of sugar composition and quantity in each sample by HPLC (Shimadzu Prominence-i Series LC-2030C system, Columbia, MD). The chromatography column was previously calibrated using known concentrations of glucose, arabinose, xylose, galactose, and mannose. Ultrapure water was used as eluent at a flow rate was 0.6 mL/min and the column temperature was maintained at 80 ˚C. A refractive index detector (model RID-20A) was maintained at 35 ˚C for sugar identification. Furfural and 5-hydroxymethyl furfural were determined using a UV–Vis L-2420 (system-integrated UV–Vis) detector. A technical and biological replicate was measured for each sample.

3.5). References

1. Cheng, G.; Zhang, X.; Simmons, B.; Singh, S., Theory, practice and prospects of X-ray and neutron scattering for lignocellulosic biomass characterization: towards understanding biomass pretreatment. *Energy & Environmental Science* **2015,** *8* (2), 436-455.
2. Heller, W. T.; Urban, V. S.; Lynn, G. W.; Weiss, K. L.; O'Neill, H. M.; Pingali, S. V.; Qian, S.; Littrell, K. C.; Melnichenko, Y. B.; Buchanan, M. V.; Selby, D. L.; Wignall, G. D.; Butler, P. D.; Myles, D. A., The Bio-SANS instrument at the High Flux Isotope Reactor of Oak Ridge National Laboratory. *Journal of Applied Crystallography* **2014,** *47* (4), 1238-1246.
3. Higgins, J. S.; Benoit, H. C., *Polymers and neutron scattering*. Clarendon Press: Oxford, 2009.
4. Roe, R. J., *Methods of X-ray and neutron scattering in polymer science*. Oxford University Press: New York, 2000.
5. Schnablegger, H.; Singh, Y., *The SAXS Guide Getting acquinted with the principles*. 4th ed.; Anton Paar GmbH: 2017.
6. Guinier, A.; Fournet, G.; Yudowitch, K. L., Small-angle scattering of X-rays. **1955**.
7. Cullity, B. D.; Stock, S. R.; Pearson India Education, S., *Elements of X-ray diffraction*. Pearson India Education Services: [Miejsce nieznane], 2015.
8. Beaucage, G., Approximations Leading to a Unified Exponential/Power-Law Approach to Small-Angle Scattering. *Journal of Applied Crystallography* **1995,** *28* (6), 717-728.
9. Park, S.; Baker, J. O.; Himmel, M. E.; Parilla, P. A.; Johnson, D. K., Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnology for Biofuels* **2010,** *3* (1), 10.
10. Garvey, C. J.; Parker, I. H.; Simon, G. P., On the Interpretation of X-Ray Diffraction Powder Patterns in Terms of the Nanostructure of Cellulose I Fibres. *Macromolecular Chemistry and Physics* **2005,** *206* (15), 1568-1575.
11. Hermans, P. H.; Heikens, D.; Weidinger, A., A quantitative investigation on the X-ray small angle scattering of cellulose fibers. Part II. The scattering power of various cellulose fibers. *Journal of Polymer Science* **1959,** *35* (128), 145-165.
12. A. Sluiter, B. H., R. Ruiz, C. Scarlata, J. Sluiter, and D. Templeton Determination of Sugars, Byproducts, and Degradation Products in Liquid Fraction Process Samples. *Technical Report NREL/TP-510-42623* **2006**, 1-11.
13. Dumitrache, A.; Natzke, J.; Rodriguez, M.; Yee, K. L.; Thompson, O. A.; Poovaiah, C. R.; Shen, H.; Mazarei, M.; Baxter, H. L.; Fu, C. X.; Wang, Z. Y.; Biswal, A. K.; Li, G. F.; Srivastava, A. C.; Tang, Y. H.; Stewart, C. N.; Dixon, R. A.; Nelson, R. S.; Mohnen, D.; Mielenz, J.; Brown, S. D.; Davison, B. H., Transgenic switchgrass (Panicum virgatum L.) targeted for reduced recalcitrance to bioconversion: a 2-year comparative analysis of field-grown lines modified for target gene or genetic element expression. *Plant Biotechnol J* **2017,** *15* (6), 688-697.

Chapter 4). Hemicellulose-cellulose composites reveal differences in cellulose organization after dilute acid pretreatment

A version of this chapter was published in a peer-reviewed journal as follows:

Shah R., Huang S., Sawada D., Pu Y, Rodriguez M, Pingali S.V., Ragauskas A.J, Kim S., Evans B.R., Davison B.H and O’Neill H “Hemicellulose-cellulose composites reveal differences in cellulose organization after dilute acid pretreatment”, Biomacromolecules, 2019, 20, 2, 893-903

4.1). Abstract

Model hemicellulose-cellulose composites, that mimic plant cell wall polymer interactions were prepared by synthesizing deuterated bacterial cellulose in the presence of glucomannan or xyloglucan. Dilute acid pretreatment (DAP) of these materials was studied using small-angle neutron scattering, X-ray diffraction and sum frequency generation spectroscopy. The macrofibril dimensions of the pretreated cellulose alone were smaller but with similar entanglement of macrofibrillar network as native cellulose. In addition, the crystallite size dimension along the (010) plane increased. Glucomannan-cellulose underwent similar changes to cellulose, except that the macrofibrillar network was more entangled after DAP. Conversely, in xyloglucan-cellulose the macrofibril dimensions and macrofibrillar network were relatively unchanged after pretreatment but the cellulose I_{β} content was increased. Our results point to a tight interaction of xyloglucan with microfibrils while glucomannan only interacts with macrofibril surfaces. This study provides insight into roles of different hemicellulose-cellulose interactions and may help in improving pretreatment processes or engineering plants with decreased recalcitrance.

4.2). Introduction

We are highly dependent on crude oil for transportation fuel and other hydrocarbon derived products. However, this is a non-renewable source and the reservoirs of crude oil are limited making it critical to find an alternative source from which transportation fuels can be derived.[1] The carbon rich polymers such as cellulose, hemicellulose and lignin make up approximately 80% of wood and other plant biomass making it an excellent source of carbon-based fuels and bioproducts. As per the 2016 Billion-Ton report, by the year 2030 the U.S. will have approximately 1 billion dry tons per year just for energy purposes and this would be sufficient to meet over 25% of transportation energy needs for this country.[2]

Physical, biological or thermochemical pretreatment approaches are commonly used as an initial step to disrupt the complex cell wall structure.[3] This enables greater access to cellulose so it can be converted to glucose using enzymatic hydrolysis and fermentation to produce ethanol that serves as transportation fuel. Different pretreatment approaches have been developed and optimized for particular feedstocks, to increase economic competitiveness and to improve sugar yields.[4] A variety of chemicals have been used for thermochemical pretreatment that can broadly be categorized into aqueous (e.g. dilute acid, alkali, ammonium-percolation, soda-lime) or non-aqueous in nature (ethanol- water, acetone-water, ionic liquids).[4-7] Beneficial solvent interactions with biomass components, maximize desired end-product yield and is an important factor for the selection of the pretreatment approach. Understanding the changes that occur due to solvent-polymer interactions can help maximize sugar and bioproduct production.

In particular, dilute acid pretreatment (DAP) is a common industrially used pretreatment for producing biofuels such as ethanol because acid not only helps to breakdown the plant cell wall structure but also hydrolyzes the matrix polysaccharides by acting on the glycosidic linkage.[4] In this pretreatment approach, dilute acid (0.2 to 2.5% w/w) is added to the feedstock which is then heated to 120 - 200 ˚C with constant mixing in a closed reactor.[7] It has been extensively studied and the reactions conditions have been optimized in order to maximize sugar yield from a variety of feedstocks.[8-11] The structural changes that occur in the cell wall after DAP are well documented and include coalescence of cellulose microfibrils, hemicellulose dissolution into the solvent, and formation of lignin aggregates at the surface of the cellulose microfibrils.[12-15] A major advantage of DAP is that it drastically increases the accessibility of cellulose to enzyme attack and also solubilizes hemicellulose to simple sugars.[3] However, the formation of inhibitory products from subsequent hemicellulose degradation such as furfural and hydromethyl furfural (5-HMF), lignin aggregation and corrosion of the reactor vessel are considered disadvantages that hinder sugar yield and downstream fermentation.[4, 12, 16] There has been some effort put into mitigating disadvantages of DAP such as feedstock selection and modulating reaction conditions such as acid concentration, temperature and pretreatment time.[3, 5, 8] The presence of lignin aggregates formed at the surface of cellulose microfibrils during DAP has been proposed to inhibit cell wall hydrolysis by preventing enzyme access to cellulose and also to unproductively bind enzymes.[17-18]

Our understanding of polymer-polymer interactions in the plant cell wall is limited making it difficult to fully understand forces that impact pretreatment efficiency at the molecular-level in natural complex multi-polymeric plant cell walls. An alternative approach to study polymer-polymer interactions between biomass polymers is to use simplified model systems as mimics of the plant cell wall. This can enable us to better understand how individual polymers interact with pretreatment solvents and to study their fate during pretreatment. Plant cell wall components have been extracted and studied individually and in combination with each other to learn about their structures, properties and interactions. For instance, native celluloses from bacteria, algae and tunicates, and their crystalline allomorphs, have been studied to determine how the glucan chain packing is altered as a consequence of different thermochemical pretreatment processes.[19-20] The interactions of hemicellulose and cellulose have also been widely studied. The binding mechanism and binding affinity of xyloglucan and cellulose was studied by coating different concentrations of xyloglucan on a layer of deposited cellulose and analyzed by reflectivity[21] and a quartz crystal microbalance[22] to show that there is a relation between the concentration of xyloglucan and its conformation on cellulose surface. In addition, the conformation of xylan from birch kraft pulp or pine sulfite was shown to transition from a threefold helical screw in bulk solution to a two-fold helix when bound to cellulose.[23] The examples described above demonstrate that even though the biomass copolymers are not in their native plant environment they retain specific characteristics that are useful in understanding *in planta* processes.

One of the disadvantages of the approaches described above for studying matrix copolymer interactions with cellulose is that it is only possible to determine how these polymers interact with intact microfibrils and it is not possible to study how these polymers may influence interactions with cellulose microfibrils as they are formed. Bacterial cellulose, produced by the genus *Acetobacter xylinus*, has been proposed as a surrogate material to study the formation of cellulose microfibrils and their interactions with other materials.[24-29] This type of cellulose is secreted by bacteria as they grow and forms a crystalline cellulose network at the surface of the bacterial culture.[29] The cellulose microfibrils are formed from the self-assembly of the cellulose chains synthesized by the cellulose synthesis terminal complex ,that in turn form microfibril bundles, termed macrofibrils.[25, 30-31] The mechanism of assembly of the cellulose chains into microfibrils and finally to macrofibrils, as determined using transmission electron microscopy[32-34] and other analytical techniques,[24-26, 31, 35] is similar to what occurs in plants but there are differences in the microfibril structure and hierarchical organization between the two materials.

Several studies have investigated the interactions of cellulose and hemicellulose by incorporating the hemicellulose into the growing cellulose network.[24-29, 31, 35-37] These studies have identified differences in the affinity of different hemicelluloses for cellulose and how they interact with the growing cellulose network. For instance, X-ray diffraction analysis (XRD) showed that beech xylan, spruce galactoglucomannan and xyloglucan from tamarind extract reduces the crystal size and crystallinity of cellulose whereas wheat arabinoxylan did not.[25, 31, 38-39] In addition, the presence of beech xylan, spruce galactoglucomannan or xyloglucan in growing cultures of bacterial cellulose changes the major crystalline form of cellulose from cellulose I_α, the dominant allomorph in bacterial cellulose, partially to cellulose I_β, that is mainly found in plants.[40-41]

In this study, we studied the effect of DAP on hemicellulose-cellulose composite materials to gain insight into the molecular-level interactions of these polymers and to understand the changes that occur in the hierarchical structure of cellulose due to the presence of hemicelluloses. The composites were prepared by synthesizing deuterated bacterial cellulose from *Acetobacter xylinus* subsp. *sucrofermentans* in presence of glucomannan or xyloglucan dissolved in the growth media.[42] Contrast matching small-angle neutron scattering (SANS) was performed in 45% D_2O solvent, the contrast match point experimentally determined for protiated hemicellulose, making it possible to obtain information about the nanoscale changes in cellulose structure. The native and pretreated cellulose and hemicellulose-cellulose composites were also studied with XRD and sum frequency generation spectroscopy (SFG) in order obtain atomic length scale structural information. The native cellulose and the hemicellulose-cellulose composites show changes in the both the arrangement of microfibrils and also the cellulose macrofibrillar network after DAP. Our data show that glucomannan most likely interacts at the surface of the macrofibrils because it does not change the crystalline form of the cellulose and the macrofibrils collapse as a result of DAP, similar to native cellulose. In contrast, xyloglucan likely interacts directly with the cellulose microfibrils as they are formed. This is supported by increased cellulose I_β content and the size of the microfibril remaining unchanged after DAP. This study provides insight into hemicellulose-cellulose interactions that can be useful for developing better extraction processes for cell wall polymers and development of cellulosic bioproducts.

4.3). Material and methods

4.3.1). Preparation of cellulose and hemicellulose-cellulose composites

The bacterial strain *Acetobacter xylinus* subsp. *sucrofermentans* (ATCC 700178) was obtained from the American Type Culture Collection (Manassas, Virginia, USA). Deuterium oxide (D_2O) was purchased from Cambridge Isotope Laboratories (Andover, Massachusetts, USA). Xyloglucan (from tamarind extract) and glucomannan (from Konjac, high viscosity) were purchased from Megazyme (Co. Wicklow, Ireland). The molecular weights of two hemicellulose samples were determined using gel permeation chromatography after acetylation.[43] The xyloglucan has high molecular weights (i.e., M_w = 530,350 g/mol, M_n = 138,540 g/mol), while the glucomannan has much lower molecular weights (i.e., M_w = 75,570 g/mol, M_n = 41,680 g/mol). NMR analysis showed that the xyloglucan polymer is formed with β (1-4)-linked glucan backbone and is branched. Galactose, xylose and fucose subunits are the prominent sugars in the branched regions of xyloglucan. The glucomannan is mainly composed of β-(1-4) linked D-mannosyl and D-glucosyl units in the ratio of ~1:0.9 with a minor presence of O-acetyl groups.

The deuterated bacterial cellulose was prepared by using hydrogenated glycerol as carbon source, as previously described.[42, 44]In order to prepare hemicellulose-cellulose composites, 1% (w/v) xyloglucan or 0.5% (w/v) glucomannan stock solutions were prepared in D_2O and diluted to a final concentration of 0.5% (w/v) and 0.25% (w/v), respectively, in the deuterated growth media.[42, 44] As before, hydrogenated glycerol was used as the carbon source. A D_2O adapted *A. xylinus* sucrofermentans inoculum was added to the media (1:10 dilution) and the cultures were grown for 5 days at 26˚C. The pellicle formed at the air-liquid interface was removed and washed rigorously and repeatedly with water at 4˚C to remove the growth media and bacterial debris.[29] The pellicles were then frozen at -80˚C for 2 h before grinding (6 x 30 sec pulses, 10 secs between pulses) to form a slurry using a blender (Waring Inc, Texas, USA). The mass fraction of the hemicellulose-cellulose in the slurry, expressed as weight percent (wt %), was determined by freeze-drying a known volume of the slurry (in triplicate).

4.3.2). Quantitative estimation of sugar in cellulose and composites

The monosaccharide composition of cellulose and the hemicellulose-cellulose composites was determined based on a previously reported procedure.[45] Samples (~50 mg each) were freeze-dried to constant weight (typically 24 h) prior to analysis. Details of the compositional analysis by acid hydrolysis and HPLC are presented in the Supporting Information file. Technical and biological replicate was measured for each sample. The data presented in Figure 1 is the average of these measurements.

4.3.3). Dilute acid pretreatment (DAP) Protocol

A known volume, typically 15 – 20 ml, of a 10 mg/ml cellulose slurry in 1% (w/v) sulfuric acid was added to a 30 mL capacity pressure vessel of a Parr reactor system (Multi-Reactor, 5000-

series). After sealing, the reactor was heated to 170 ˚C at a ramp up rate of 5˚C/min and maintained at 170 ˚C for 5 mins before cooling to room temperature. The pressure recorded at 170 ˚C was between 9 and 11 atm for each reaction.

4.3.4). X-ray diffraction data collection and analysis

The cellulose and hemicellulose-cellulose composites, before and after pretreatment, were ground and freeze dried before data collection. XRD data were collected in transmission mode on a R-Axis IV detector from a rotating anode X-ray generator, MicroMAX-007HF (Rigaku) operated at 30mA and 40kV, using CuKα radiation (λ = 1.5418 Å). Exposure time for each measurement was 480 sec. The data were converted to a 1D peak profile by using Nika package[46] implemented in the Igor software package. A linear fitting was performed to remove the instrumental background. Subsequently, background intensity was estimated using the smoothing method[47] with the Savitzky-Golay filter in the 2θ range from 5 to 60 ° for each diffraction profile. The window size and polynomial order for the Savitzky-Golay filter was set to 35 and 1, respectively and the fitting procedure was repeated 100 times. The crystallinity index was calculated with the peak deconvolution approach by using the following equation:

$$\%\,Crystallinity = \frac{(total\ scattering\ area - amorphous\ area)}{total\ scattering\ area} \times 100 \tag{1}$$

where ‘total scattering area’ and ‘amorphous area’ are the integrated areas under the curves in the 2θ range from 5° to 60°. It was also calculated separately using the peak height method by taking the ratio of peak height between (110) and 2θ = 18°.

The 1D curve was normalized by dividing the intensity at all points of 1D diffraction pattern by the total scattering area and the peaks were fit using four pseudo-Voigt functions using fitting software Fityk[48] to obtain the peak position, peak height and the full width half maximum (FWHM) for each sample (Figure S1). In the cellulose-only samples, it was assumed that the major contributor of peak broadening was due to the crystallite size in all directions.[49] The instrumental broadening was calculated by analyzing the diffraction pattern of silicon (NIST standard: SRM640e) and subtracting it from the FWHM value. The crystallite size was determined based on the Scherrer equation given as:

$$t = (0.9 \times \lambda) \div (\beta \times \cos\theta) \tag{2}$$

where, t is the size of the crystallite, λ is the wavelength of the incident X-rays, β and 2θ are the values for FWHM and peak position, respectively, obtained as described above. FWHM for peaks at (100), (010) and (110) were used to determine crystallite size in the particular dimension. Two biological replicates and three technical replicates for each sample were studied.

4.3.5). Sum frequency generation (SFG) spectroscopy data collection and analysis

The native and pretreated cellulose composites were freeze-dried before the SFG analysis. The SFG experiment was performed in the reflection mode using the broadband SFG table top spectroscopic system. All SFG spectra were collected from 2700 cm^{-1} to 3700 cm^{-1} and normalized with the power profile of the incidence infrared light. Each spectrum reported in Figure 2 represents an average of three replicates from different locations on the sample. To compare the relative peak intensities in the OH region (3200-3700 cm^{-1}) between the native and pretreated samples (cellulose, xyloglucan-cellulose and glucomannan-cellulose), SFG spectra are normalized using peak intensity observed at 3320 cm^{-1} within each pair of samples. To estimate the fraction of cellulose I$_\alpha$, the broad peak in the OH region was deconvoluted into a total of 6 peaks centered at 3240, 3270, 3296, 3330, 3370, and 3450 cm^{-1} through peak fitting with a Lorentzian peak shape.[50] During the fitting, peak positions were allowed to vary in the range of 3236−3244, 3270−3275, 3296−3305, 3328−3335, and 3369−3379 cm^{-1}; their full width at half-maximum (FWHM) values were relatively constant at ~20 cm^{-1}. The position and width of the broad shoulder peak at ∼3450 cm^{-1} were allowed to relax during the fitting. The fitted peak areas at ∼3240 and ∼3270 cm^{-1} were used to calculate the fraction of I_α cellulose given by $A_{3240}/(A_{3240} + A_{3270})$ from each sample.

4.3.6). Small-angle neutron scattering (SANS)

SANS measurements were performed using the Bio-SANS instrument located at the High Flux Isotope Reactor facility in Oak Ridge National Laboratory.[51] A large dynamic *Q*-range, spanning $0.003 < Q$ (Å^{-1}) < 0.4 was accessed in a single configurational setting using 6 Å neutrons and a relative wavelength spread ($\Delta\lambda/\lambda$) of 15%. The main and wing detector arrays were positioned at 15.5 m and 1.13 m from the sample position, and the wing detector array was positioned at an angle of 1.4 ° with respect to the direct beam.[52] The samples were initially exchanged into 45%/55% (v/v) D_2O/H_2O solvent (neutral pH) over 24 h which included at least three exchanges into fresh solvent to ensure complete D/H exchange. Prior to measurement, the samples were further exchanged into the DAP solvent, 45%/55% (v/v) D_2O/H_2O containing 1% (w/v) H_2SO_4 and placed in 1-mm-thick quartz cells (Hellma Model# 124-QS 1.0 mm circular cell). The scattering intensity profile, $I(Q)$ versus Q, for each sample was obtained by azimuthally averaging the processed 2D images which were normalized to incident beam monitor counts, corrected for detector dark current, pixel sensitivity and scattering from backgrounds such as solvent and quartz cell.

4.3.6.1). SANS data fitting and analysis

The SANS data were fit to the multi-level Unified Fit model implemented in the IRENA package of Igor Pro software by Wavemetrics.[46, 53] The SANS intensity profile is a summation of individual levels (*i*) with each level modeled as the sum of an exponential and a power law behavior-

$$\boldsymbol{I(q) = \sum_{i} I_i(q) = G_i e^{-\frac{q^2 R_{g_i}^2}{3}} + e^{-\frac{q^2 R_{gCO_{i-1}}^2}{3}} \times B_i \left\{ \frac{\left(erf\left(\frac{qR_{g_i}}{\sqrt{6}} \right) \right)^3}{q} \right\}^{P_i} + C} \qquad (3)$$

where R_{g_i} is the radius of gyration of the particle of the *i*th structural level, P_i is the power-law exponent of the *i*th structural level, G_i is the scalar for the Guinier function of the *i*th structural level, and B_i is the scalar for the power-law function of the *i*th structural level. Additionally, $R_{gCO_{i-1}}$ is the cut-off length scale of the power-law behavior of the *i*th structural level and *C* is the *q*-independent constant background intensity.

4.4). Results

4.4.1). Quantitative estimation of sugar in cellulose and composites

The monosaccharide composition of native and pretreated cellulose, xyloglucan-cellulose and glucomannan-cellulose composites, shown in Figure 4.1 were determined using a published procedure.[45] Monosaccharide analysis shows that native cellulose is solely composed of glucose and the hemicellulose-cellulose composites synthesized in presence of either xyloglucan or glucomannan contained ~12% xylose and ~25% mannose, respectively.

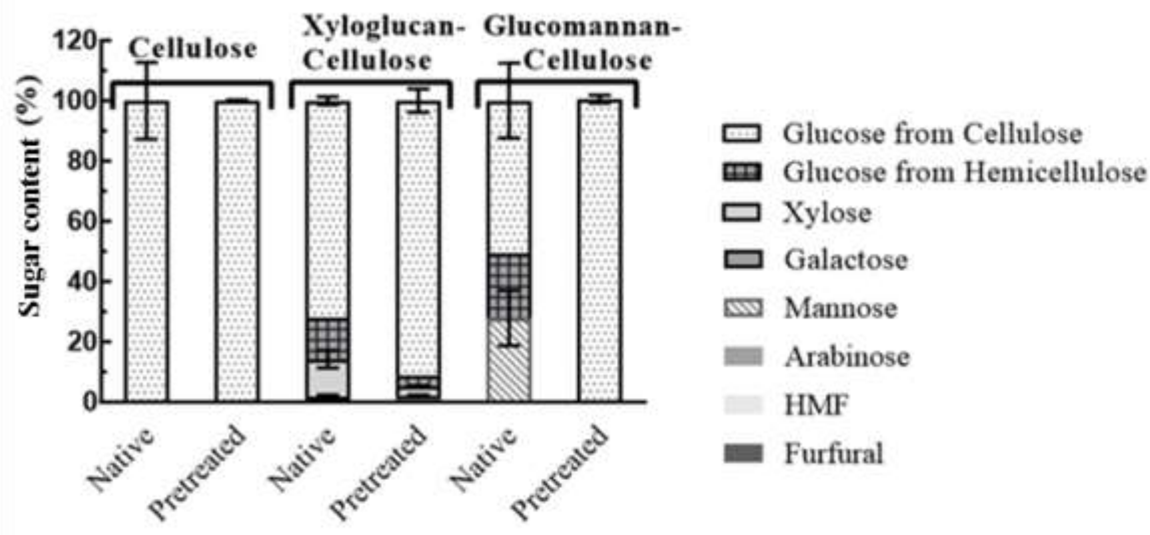


Figure 4.1: The monosaccharide composition of cellulose, xyloglucan-cellulose and glucomannan-cellulose composites before (native) and after (pretreated) dilute acid pretreatment.

After DAP, the monosaccharide composition of cellulose remained unchanged, while in the xyloglucan- and glucomannan-cellulose composites ~5% xylose and <1% of mannose remained associated with the cellulose. In summary, a significant incorporation of both xyloglucan and glucomannan was detected in the native hemicellulose-cellulose composites and although the DAP protocol completely removed glucomannan, a small amount of xyloglucan was found to remain associated with the cellulose.

4.4.2). X-ray diffraction analysis

The X-ray diffraction (XRD) profiles of native and pretreated bacterial cellulose, xyloglucan-cellulose and glucomannan-cellulose are shown in Figure 4.2. The peak positions of native bacterial cellulose are represented by planes (100), (010) and (110), typical of cellulose with a high cellulose I_α content.[41]The peak intensity of the (100) peak is double the (010) peak indicating that the crystallites in the native cellulose macrofibril are highly oriented.[26, 37, 39] However, after pretreatment the (100) and (010) peaks are of equal intensity, which occurs when there is no preferential orientation of the crystallites. In contrast, the (100) and (010) peaks of the hemicellulose-cellulose composites are of equal intensity before and after pretreatment indicating that there is no preferential orientation of crystallites in these materials.The diffraction peaks of the hemicellulose-cellulose composites, especially in the (100) and (010) planes are significantly broader than for cellulose alone (Figure 4.2) which is consistent with previous reports.[25, 37, 39]This indicates an increase in the amorphous content of these materials, and is consistent with the presence of the hemicellulose. The crystalline content of the cellulose and the hemicellulose-cellulose composites was calculated using both the peak height method and the peak deconvolution method.

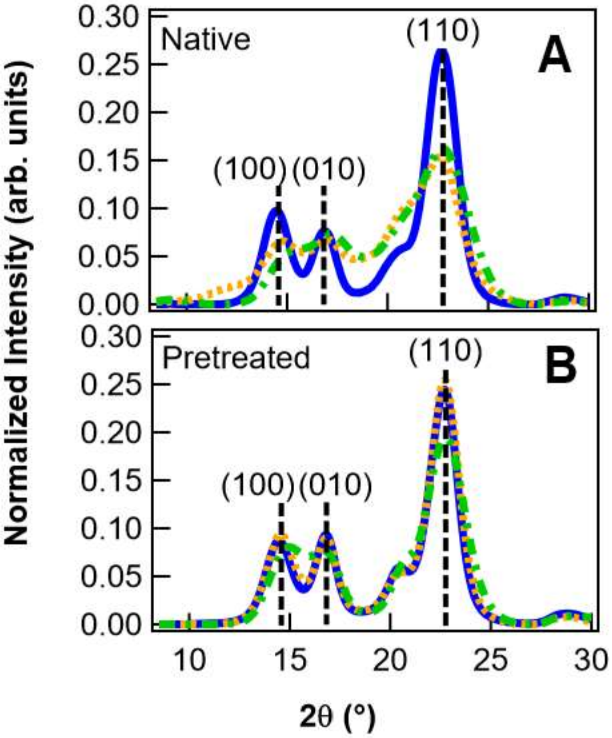


Figure 4.2: X-ray diffraction data of the native and pretreated samples.

Cellulose (blue), xyloglucan-cellulose (dashed green) and glucomannan-cellulose (dashed yellow). The reflections (100), (010) and (110) are indicated. All curves have been processed as described in Materials and Methods.

The values obtained for the samples by each method, for the native and pretreated material, showed similar trends but the absolute values for the crystalline content cellulose and the composites were higher using the peak height method (Table 4.1). We used the peak convolution method for further analysis because it takes into account the peak width change.[54] In the case of the hemicellulose-cellulose composites, sugar analysis showed that the majority of the amorphous hemicellulose is removed after DAP and this is reflected in the decrease in the amorphous part of the peak area in the XRD profiles for these samples. However, the absolute values for the crystalline content of the treated composites are lower than native and pretreated bacterial cellulose suggesting that there is a lower amount crystalline cellulose in the native hemicellulose-cellulose composites compared to cellulose alone.

Although the overall crystallinity of cellulose remained unchanged after pretreatment (Table 4.1), the dimension in the (010) plane significantly increased. In the case of glucomannan-cellulose, the increase in the crystallite dimensions was also observed after pretreatment and it is consistent with the removal of glucomannan during pretreatment. Interestingly, the values are similar to pretreated cellulose suggesting that the structural changes in the cellulose microfibril were the same for these samples. The crystallite sizes decreased in the presence of xyloglucan along all three planes ((100), (010) and (110)) compared to cellulose (Table 4.2). The crystallite sizes of xyloglucan-cellulose are the same before and after pretreatment as indicated by p-value (0.31) > 0.05 and smaller than pretreated cellulose and glucomannan-cellulose based on a p-value (0.045) < 0.05. This suggests that there is a difference in how xyloglucan and glucomannan interact with the cellulose microfibril.

Table 4.1: Estimation of cellulose crystallinity native and pretreated samples.

Sample	Crystallinity$^+$ (%)	
	Peak deconvolution	Peak height
	Native	
Cellulose	69 ± 4	95 ± 5
Xyloglucan-cellulose	37 ± 2	69 ± 8
Glucomannan-Cellulose	48 ± 5	71 ± 8
	Pretreated	
Cellulose	71 ± 3	95 ± 4
Xyloglucan-cellulose	53 ± 2	94 ± 3
Glucomannan-Cellulose	63 ± 3	95 ± 3

$^+$Errors bars were generated by averaging three biological replicates and shown as ± standard deviation

Table 4.2: Determination of crystallite size of cellulose in native and pretreated samples

Sample	Crystallite Size^{+*} *t*, (Å)		
	(110)	(010)	(100)
	Native		
Cellulose	46 ± 2	53± 3	51 ± 3
Xyloglucan-cellulose	35 ± 7	37 ± 8	43 ± 1
Glucomannan-Cellulose	49 ± 2	53 ± 1	46 ± 1
	Pretreated		
Cellulose	51 ± 2	64 ± 1	48 ± 2
Xyloglucan-cellulose	38 ± 3	43 ± 6	42 ± 2
Glucomannan-Cellulose	47 ± 1	60 ± 2	48.0 ± 1

+Error bars were generated by averaging three biological replicates and shown as ± standard deviation; *Instrumental broadening estimated using NIST standard of silicon was subtracted from FWHM.

The peak positions in the X-ray diffractograms for cellulose and hemicellulose-cellulose composites gives insight into the packing of the glucan chains in the microfibril (Table 4.3). Native bacterial cellulose has peak positions similar to those reported previously.[41] With the exception of a slight shift in (010) plane from 16.86˚ to 16.77˚, the peak positions remain unchanged after DAP. In the case of the hemicellulose-cellulose composites, the peak positions were difficult to determine accurately due to the large contribution of amorphous scattering from the hemicellulose in the samples (Table 4.3). Removal of the hemicellulose component of the composites during pretreatment enabled the peak positions to be determined with a higher degree of confidence. The peak positions determined for glucomannan-cellulose composite were similar to native bacterial cellulose after pretreatment indicating that glucomannan does not interfere with glucan chain packing in the composite. Conversely, the peak positions in xyloglucan-cellulose were shifted compared to native cellulose, suggesting that xyloglucan interferes with the glucan chain packing in the cellulose microfibril during formation. Prior studies have reported similar results for xyloglucan-cellulose composites[25, 41, 55] that were interpreted as a change in the cellulose allomorph from cellulose I_α to I_β.

4.4.3). Sum frequency generation spectroscopy

Sum frequency generation spectroscopy (SFG) was used as a complementary technique to X-ray diffraction to measure the crystalline content of the cellulose and hemicellulose composites. Figure 4.3 shows the averaged and normalized SFG spectra for native and pretreated cellulose and composites. The information in the spectral regions for SFG analysis was in the alkyl stretching (2700−3050 cm^{-1}) and hydroxyl stretching (3150−3700 cm^{-1}) regions. The spectral features and intensities of the native cellulose are almost identical before and after pretreatment (Figure 4.3), suggesting that the cellulose structure is not affected by DAP.

Table 4.3: The peak positions of the X-ray diffraction data determined by fitting the overall scattering intensity in native and pretreated samples.

Sample	Peak Position$^+$, 2θ (°)		
	(110)	(010)	(100)
	Native		
Cellulose	22.72 ± 0.02	16.86 ± 0.02	14.47 ± 0.01
Xyloglucan-cellulose	22.6 ± 0.1	17.0 ± 0.2	14.7 ± 0.1
Glucomannan-Cellulose	22.65 ± 0.01	16.91 ± 0.01	14.66 ± 0.02
	Pretreated		
Cellulose	22.68 ± 0.02	16.77 ± 0.01	14.47 ± 0.02
Xyloglucan-cellulose	22.57 ± 0.01	16.66 ± 0.03	14.70 ± 0.03
Glucomannan-Cellulose	22.63 ± 0.07	16.75 ± 0.05	14.55 ± 0.05

+Error bars were generated by averaging three biological replicates and shown as ± standard deviation.

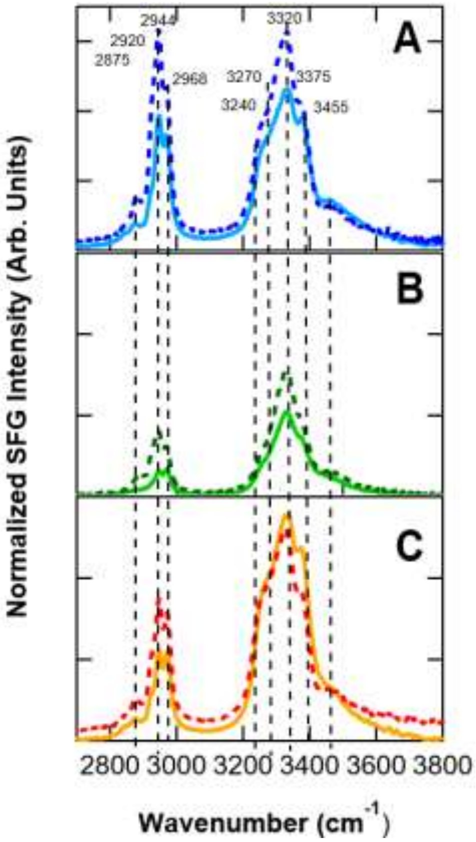


Figure 4.3: The SFG intensity for the native cellulose (blue), xyloglucan-cellulose (green) and glucomannan-cellulose (orange) composites and pretreated counterparts (dashed line) were recorded versus wavenumber. Each curve is an average of three technical replicates collected from different locations in the sample. The curves were normalized to the peak intensity observed at 3320 cm^{-1} within each pair of samples

This is also observed for the xyloglucan-cellulose and glucomannan-cellulose composites indicating that the crystallinity in the native and pretreated material is similar.

The peaks at 3240 and 3270 cm^{-1} are characteristic of the presence of cellulose I_α and I_β, respectively. By deconvoluting these two peaks the relative amount of cellulose I_α and I_β allomorphs in cellulose and composites was estimated (Figure S2). As previously observed,[37] the incorporation of xyloglucan increases the amount of cellulose I_β in the xyloglucan-cellulose composites. Whereas for the glucomannan-cellulose no significant change in the ratio of the two allomorphs were observed. After pretreatment, the I_α/I_β ratio remains unchanged in each hemicellulose-cellulose composite. This suggests that xyloglucan can interact with glucan chain and changes the allomorph of the crystallite in the microfibril. On the other hand, glucomannan, does not appear to change the crystallinity or the ratio or I_α/I_β suggesting that it does not directly interact with the glucan chain during formation of the cellulose microfibrils.

4.4.4) Small angle neutron scattering analysis

SANS was used to study how the nanoscale structural properties and organization of cellulose is changed in the hemicellulose-cellulose composites during DAP. SANS is ideally suited for this type of study because it probes a large range of length scales from 1 – 500 nm.[51]. In addition, the difference in the neutron scattering cross-section of hydrogen and deuterium can be used to selectively highlight the structural characteristics of individual components in a complex system by varying the H_2O/D_2O ratio of the solvent[51].

As described in the Materials and Methods section, the hemicellulose-cellulose composites were prepared by adding hemicellulose to a culture of *A. xylinus* subsp. *sucrofermentans* growing in the D_2O based media. This resulted in the synthesis of partially deuterated cellulose and protiated hemicellulose composite. We have previously determined the level of deuterium incorporation in cellulose under these conditions.[44] The neutron contrast match point for glucomannan was obtained from scattering profiles recorded for glucomannan in H_2O/D_2O ratios between 0 and 100%, as described previously.[42] The scattering intensity increased with decreasing Q values over the measured Q range indicating that the glucomannan was aggregated under the conditions measured (see Figure S3). Therefore, we determined the match point at three different Q values (see Figure 4.4), and in all cases, the calculated contrast match point for glucomannan was determined to be 45 ± 1 % D_2O indicating that there were no local density fluctuations in the glucomannan aggregates (Figure 4.4). The SANS data presented here were recorded in 45% D_2O to selectively highlight structural changes in cellulose when the hemicellulose scattering is contrast matched to the solvent. The SANS profiles of fully protiated hemicellulose-cellulose composites and cellulose alone were compared to their partially deuterated counterparts. No significant differences between the samples were observed indicating that D_2O does not cause any changes to the cellulose structure (see Figure S4).

The Unified fit approach was used to analyze the SANS data. This fitting approach was developed for analyzing SANS data for complex hierarchical systems and has been previously applied for studying biomass and purified biomass polymers.[12, 31, 56] For all samples, the data could be fit with 2 structural levels that covered the Q range from $0.007\ \text{Å}^{-1} > Q > 0.5\ \text{Å}^{-1}$ and $0.003 > Q > 0.007\ \text{Å}^{-1}$, respectively (Figure S5). The corresponding estimated real space distances ($2\pi/Q$) are ~15 Å - ~900 Å for structural level 1 and ~900 Å - ~2100 Å for structural level 2.

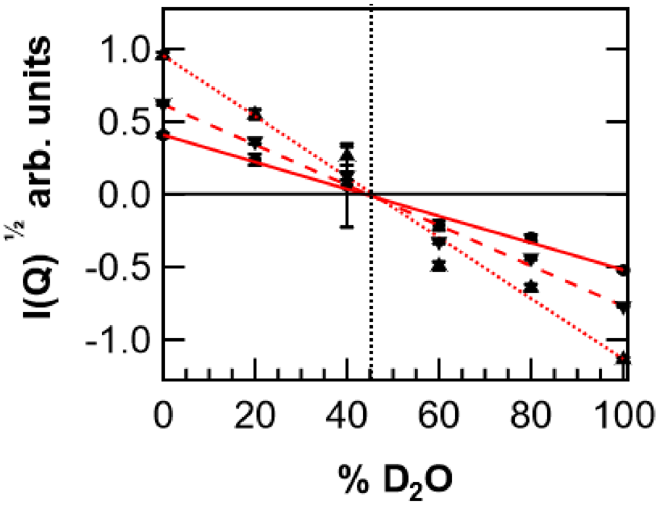


Figure 4.4: Neutron contrast match point determination for 0.5% (w/v) glucomannan. A plot of square-root intensity versus H_2O/D_2O solvent ratio for Q values at 0.005 (red dotted line), 0.007 (red dashed line) and 0.01 Å^{-1} (red line) shown here were chosen as the scattering for glucomannan as low Q does not reach a constant value. The x-value of the intersection between the straight line fit (red line) at which the $\sqrt{I} = 0$ is the neutron contrast match point

Figure 4.5 shows the scattering profiles of cellulose in native and pretreated states. The structural parameters obtained using a 2-level Unified Fit approach are presented in Table 4.4. The level-1 radius of gyration, R_g, decreased from 230 ± 32 Å to 130 ± 19 Å after pretreatment and the corresponding power-law exponent (α) increased from 2.8 ± 0.1 to 3.1 ± 0.2. The level-1 R_g value can be interpreted as the size of a cross-sectional rectangular cellulose macrofibril with a diameter (D) of ~577 Å ($D = \sqrt{\frac{16}{3}} \times R_g$).[31] This value is consistent with the reported diameter 400 – 600 Å, of a bacterial cellulose macrofibril obtained using transmission electron microscopy,[32, 34] providing confidence that this feature in the SANS curves is related to the cross-sectional diameter of a bacterial cellulose macrofibril. The decrease in the R_g value after pretreatment indicates that the macrofibril diameter is significantly decreased, as might occur due to water expulsion from the macrofibril and adjacent cellulose coalescence at high temperature and pressure that occurs during thermochemical pretreatment A similar phenomenon has been observed for the changes in cellulose microfibril packing during steam pretreatment of poplar.[45] Further, the level-1 power-law behavior in the native and pretreated material gives insight into the arrangement of the cellulose microfibrils in the macrofibril. Native cellulose exhibits an exponent of 2.8 ± 0.1 that can be interpreted as a highly entangled network of cellulose microfibrils in the macrofibril.[53] After pretreatment, the network arrangement of the cellulose microfibrils is now no longer visible, α has increased to 3.1 ± 0.2, and is interpreted as a surface fractal ($3 < \alpha < 4$ [53]) with rough interfaces with the solvent. The change in the power-law behavior from a mass-fractal to a surface fractal is consistent with the change in the macrofibril structure from a hydrated network of cellulose microfibrils to a coalesced structure, as described above.

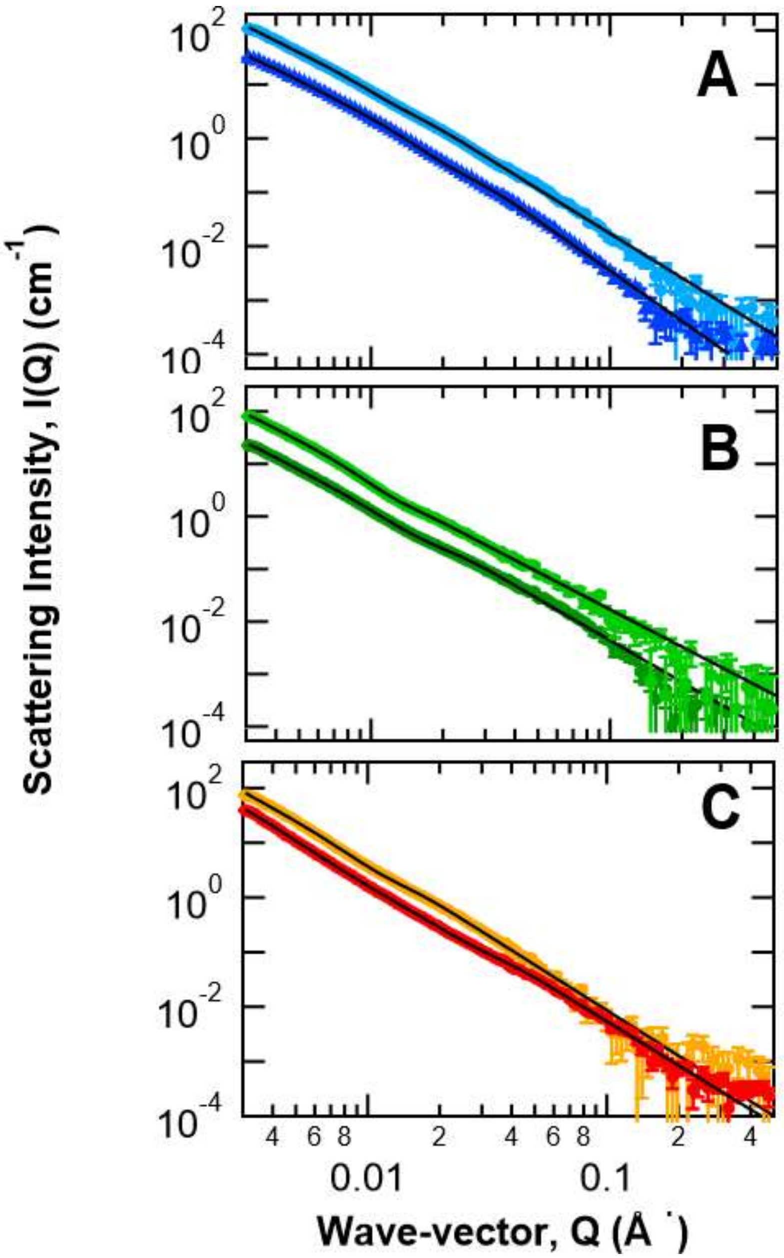


Figure 4.5: SANS profiles of native and pretreated cellulose and hemicellulose-cellulose composite samples. Cellulose (blue), xyloglucan-cellulose (green), and glucomannan-cellulose (orange). Solid red lines are the Unified Fit curves. In all cases, the pretreated cellulose and composites curves have been scaled by a factor of 5 for clarity purpose.

Table 4.4: Unified fit parameter values obtained from fitting SANS profiles of native and pretreated samples.

Structural Levels	Level - 1		Level - 2	
	R_g (Å)	α *	#R_g (Å)	α *
Samples	**Native**			
Cellulose	230 ± 32	2.8 ± 0.1	1500	2.5 ± 0.1
Xyloglucan-Cellulose	220 ± 19	2.5 ± 0.1	1500	2.1 ± 0.1
Glucomannan-Cellulose	250 ± 35	2.8 ± 0.1	1250	2.4 ± 0.2
	Pretreated			
Cellulose	130 ± 19	3.1 ± 0.2	1250	2.2 ± 0.1
Xyloglucan-Cellulose	170 ± 19	2.6 ± 0.1	1250	2.3 ± 0.1
Glucomannan-Cellulose	90 ± 29	2.9 ± 0.3	1250	2.8 ± 0.1

#Fixed (parameter not allowed to vary during fitting process). *α is the power law exponent

In the level-2 structural regime, we observe macrofibril arrangement in the bacterial cellulose network. It is not possible to accurately determine a value for level-2 R_g for both native or pretreatment sample because its size was larger than the accessible length scales (> 1000 Å). Therefore, the R_g value was fixed to a relatively large value, $R_g = 1250$ Å for the purpose of fitting. The level-2 α value, obtained from fitting the low-Q region decreased from 2.5 ± 0.1 to 2.2 ± 0.1 after DAP indicating that the network of macrofibrils in the native and pretreated cellulose are similar in native and pretreated cellulose.

The structural parameters for glucomannan-cellulose obtained before and after pretreatment using a 2-level Unified fit, showed similar trends to native cellulose. The level-1 R_g for native glucomannan-cellulose was 250 ± 35 Å and it decreased to 90 ± 29Å after pretreatment. The initial R_g value is similar to native cellulose and the decrease in the R_g can be interpreted as expulsion of water from the highly hydrated cellulose macrofibril as observed for the native cellulose. The level-1 α values before and after pretreatment are also similar to those observed for native cellulose (within error) consistent with a collapse of microfibril network structure in the macrofibril as a result of pretreatment. The level-2 structural regime was restricted to interpreting only the power-law exponent for the native cellulose sample. The level-2 α value slightly increased from 2.4 ± 0.2 in native sample to 2.8 ± 0.1 in the pretreated sample. This shows the opposite trend to that observed for native cellulose and is somewhat unexpected because the thickness of the macrofibrils was significantly reduced during pretreatment. Therefore, one would expect to form a sparse cellulose network (i.e. a decrease in α) as observed in pretreated cellulose sample.

In the case of xyloglucan-cellulose, the effect of pretreatment on the composite showed a different trend. Unlike native cellulose and glucomannan-cellulose, the level 1 R_g values remained relatively unchanged before and after pretreatment. The R_g value in native xyloglucan-cellulose was 220 ±

19Å, similar to the value obtained for the native cellulose, and decreased to 170 ± 19 Å in pretreated xyloglucan-cellulose, a value significantly larger than that obtained for both pretreated cellulose and glucomannan-cellulose composite. The level-1 α values were 2.5 ± 0.1 and 2.6 ± 0.1 for native and pretreated xyloglucan-cellulose composite, respectively. These values indicate that the arrangement of the cellulose microfibrils in the macrofibril were unchanged (within error) due to the presence of xyloglucan suggesting a different mode of interaction with the cellulose than glucomannan. The level-2 structural parameters also showed some differences compared to the native cellulose and glucomannan-cellulose composites. The level-2 α values were 2.1 ± 0.1 and 2.3 ± 0.1 in native and pretreated xyloglucan-cellulose composite, respectively. The cellulose macrofibrils in xyloglucan-cellulose assemble as a sparse or loosely entangled network (α = ~2.1) compared to the dense highly entangled network formed by glucomannan cellulose. This value is only slightly changed for xyloglucan-cellulose after pretreatment indicating that xyloglucan does not interfere with inter-macrofibril interactions and consequently, the overall structure of the pretreated cellulose network is unchanged.

4.5). Discussion

Our overall aim was to investigate the nature of cellulose interactions with hemicellulose and to examine in detail the structural changes individual hemicellulose types cause on cellulose structure during pretreatment. In previous work, we (and others) have studied structural changes in whole biomass during dilute acid pretreatment and quantified changes to the component biopolymers that comprise the cell wall.[9, 11, 13-14, 18, 57-58] Although it is commonly understood that the hemicellulose is dissolved and depolymerized by the pretreatment solvent during this process, the molecular-level changes that occur due to removal of hemicellulose are poorly understood. This is difficult to study with intact biomass samples due to the inherent complexity of the material and the challenge of deconvoluting the contributions of individual components in the hierarchical laminate formed by the plant cell wall. Bacterial cellulose has been proposed as a model cellulose that can be used as mimic to understand the plant cell wall. In previous work, hemicelluloses have been incorporated into bacterial cellulose and changes in the hierarchical structural assembly due to specific hemicellulose-cellulose interactions have been reported. Martinez-Sanz *et al.* studied the interaction of different kinds of hemicellulose with deuterated bacterial cellulose using SAXS,[25] SANS,[28] and USANS,[26] and showed that arabinoxylan deposits on the cellulose macrofibril surface while xyloglucan and mixed linked glucans interact with microfibrils present in the cellulose macrofibril. Penttila *et al.* performed a similar study using secondary plant cell wall hemicellulose like spruce xylan and galactoglucomannans and showed they both induce changes in the microfibrillar structure of the cellulose macrofibril.[31]

In our study, XRD, SFG and SANS analysis of the hemicellulose-cellulose composites, highlighted differences in cellulose structure and morphology due to DAP. These differences were related to glucomannan and xyloglucan interactions with the cellulose network as it is formed. Analysis of the structural changes in native cellulose after pretreatment showed that the cross-sectional diameter of the cellulose macrofibril and surface morphology, and the internal microfibril structure of cellulose changes to form a tightly packed cellulose macrofibril. This can be related to expulsion of water trapped between cellulose microfibrils[59] during DAP resulting in the microfibrils interacting with each other to form a tightly packed cellulose fiber. This is consistent with coalescence of the cellulose microfibrils due to water expulsion previously observed during steam explosion of poplar wood.[57] Although the XRD data for glucomannan-cellulose composite indicated a significant increase in cellulose crystallinity after pretreatment, SFG data showed that apparent increase in cellulose crystallinity was primarily due to the removal of amorphous hemicellulose, rather than changes to the microfibrils themselves. This indicates that glucomannan does not interfere with the assembly of the microfibrils by inhibiting crystallization of the cellulose chains. The nano-scale structural changes, as determined using SANS, were similar to the native cellulose sample. A significant decrease in the cross-sectional diameter of the cellulose microfibril was observed after DAP, which can be related to the expulsion of water during pretreatment, as described for cellulose above. However, the network of macrofibrils in the glucomannan-cellulose composite become more entangled after pretreatment compared to cellulose in the native and pretreated state. This allows us to speculate that glucomannan may accumulate between intersecting macrofibrils in the cellulose matrix preventing interactions so that the macrofibrils can be more easily rearranged during the pretreatment process.

Unlike glucomannan, xyloglucan does change the structure and organization of the microfibrils. Both XRD and SFG analyses show that the cellulose crystalline allomorph transitions from I_α to I_β and these structural changes persists after pretreatment. In addition, the crystallite sizes in all three dimensions are smaller than cellulose and these do not change after pretreatment. The smaller crystallite size in xyloglucan-cellulose composites have been previously reported [25, 38] and it has been inferred that xyloglucan interacts with the bacterial cellulose microfibril during its formation. The nanoscale changes in the xyloglucan-cellulose composite are also different to those observed for native cellulose and glucomannan cellulose. The cross-sectional diameter of the xyloglucan-cellulose macrofibril is relatively unchanged after pretreatment. A clue to the reason for this can be obtained from analysis of the monosaccharide composition of the hemicellulose-cellulose composites. In the case of glucomannan-cellulose, only a trace amount of glucomannan remained after pretreatment as evidenced by comparing the sugar analysis of the composite before and after pretreatment. Whereas, approximately 5% of xylose remains associated with the cellulose in the xyloglucan-cellulose composite after pretreatment. Based on this observation, we can suggest that unlike glucomannan, a fraction of the xyloglucan associated with cellulose is inaccessible to acid during the pretreatment process or there is strong interaction between xyloglucan and cellulose that prevents its complete removal during the pretreatment process. This incomplete removal of xyloglucan could prevent coalescence of the microfibrils at high temperature as has been demonstrated by our SANS results. This is supported by results of several studies that have shown that xyloglucan is present between microfibrils and not just on the surface of the macrofibrils.[21, 36, 60-61]

We also note that the crystallite size along (010) plane increases for cellulose and for glucomannan-cellulose but remains unchanged in the other two planes. Conversely, the crystallite sizes remain unchanged for xyloglucan-cellulose after DAP pretreatment. We can infer that the cellulose microfibrils coalesce along this face of the microfibrils in cellulose and glucomannan cellulose. A previous study investigated the adsorption of xyloglucan on cellulose surfaces showed that xyloglucan is adsorbed on the cellulose surface by an entropy driven process in which water is removed from the cellulose surface in order for xyloglucan to bind to it. [22] Both the (100) and (010) crystalline planes of cellulose are hydrophilic having a propensity of -OH groups projecting towards the solvent whereas the (110) is hydrophobic with the glucose planar structure is along that plane. The increase in the crystallite size is observed along the (010) plane, implying that the higher density of -OH groups in the (010) plane make it more hydrophilic and thus more favorable for xyloglucan adsorption on to this surface. It is therefore likely that xyloglucan binds to the (010) plane of the crystallite preventing coalescence of neighboring cellulose microfibrils after pretreatment.

4.6). Conclusions

The effect of DAP on the structure and morphology of the bacterial cellulose by xyloglucan, a primary plant cell wall component and glucomannan, a secondary plant cell wall component was studied using SANS, XRD and SFG spectroscopy. Figure 6 summarizes the main findings of this study. We observed changes to both the packing of microfibrils in the macrofibrils and also the macrofibrillar network itself as a consequence of DAP. Importantly, the nature of these changes was dependent on the type and structures of hemicellulose present. We can propose that glucomannan most likely interacts at the surface of the macrofibrils because it does not change the crystalline form of the cellulose and the macrofibrils collapse as a result of DAP, similar to native cellulose. On the other hand, xyloglucan interacts directly with the cellulose microfibrils as they are formed. This is supported by increased crystalline I_β content in the cellulose microfibrils and the size of the microfibril remaining unchanged after DAP. This study provides new insights into the underlying mechanism that leads to changes observed in cellulose structure due to presence of and the subsequent removal of hemicellulose during DAP. It shows that understanding how hemicellulose interacts with and affects cellulose structure during different processing regimes is critical for improving cellulosic products and developing new bioproducts and biomaterials.

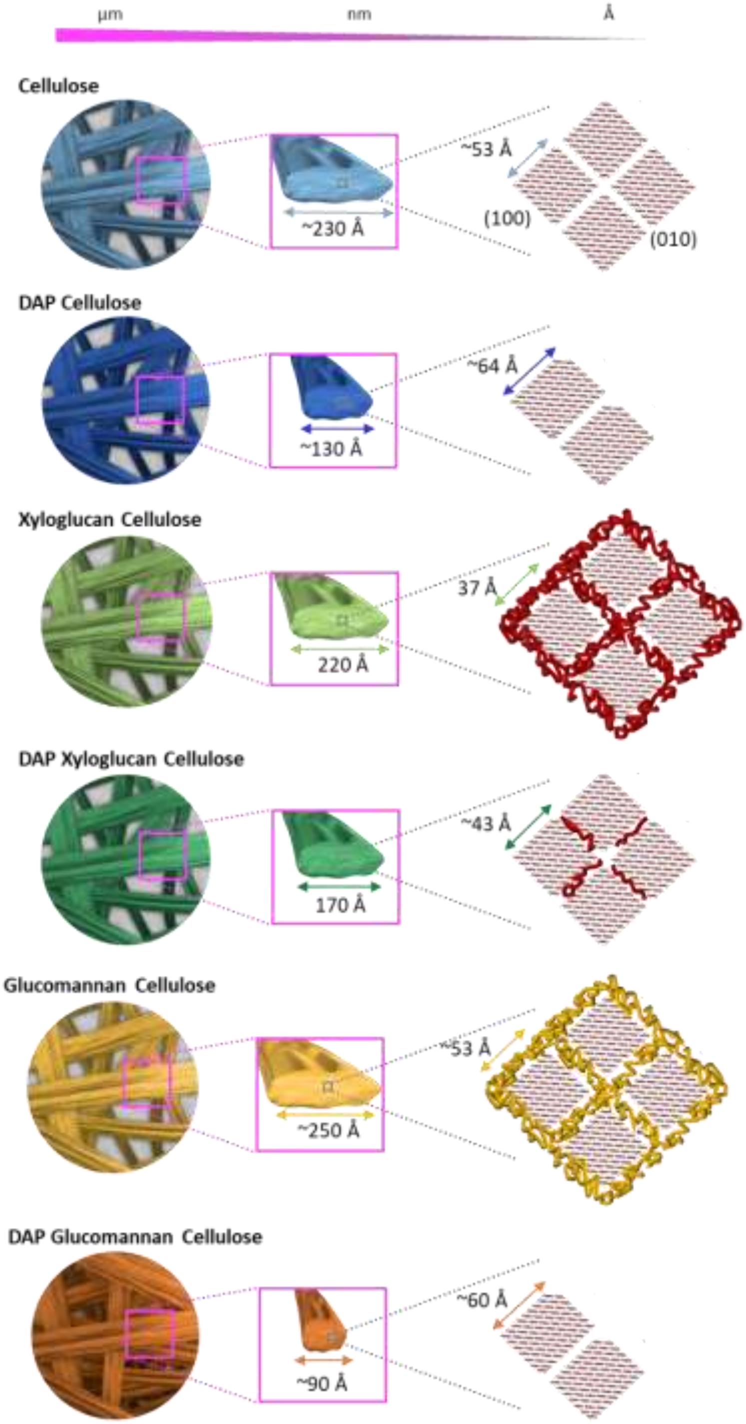


Figure 4.6: A schematic representation of the SANS and XRD data for the cellulose and hemicellulose-cellulose before and after DAP. At the nanoscale, the cellulose macrofibrils collapse after DAP and the macrofibril network that they form appears sparse. The crystallite size increases in (010) plane. For xyloglucan-cellulose the cellulose macrofibrils do not significantly change size and the crystallite size also does not increase. For glucomannan-cellulose a trend similar to that of cellulose in native and pretreated state is observed except the network of macrofibril after DAP appear more entangled

Supporting Information File:

Supplement file 1 Wide angle x-ray scattering intensity profile for native and pretreated cellulose and hemicellulose-cellulose composites with overall fit, Plot showing cellulose I_α values determined by peak deconvolution of SFG spectra for native and pretreated cellulose and composites, I(Q) vs Q plot for glucomannan dissolved in different H_2O/D_2O ratios for determination of its contrast match point, Comparison of SANS curve for hydrogenated and deuterated cellulose, Unified fit analysis of cellulose SANS profile

ACKNOWLEDGEMENTS

R. S., D. S., S. V. P., H. O'N., B.E., and B.H.D. acknowledge the support of the Genomic Science Program, Office of Biological and Environmental Research (OBER), U. S. Department of Energy, under Contract FWP ERKP752, for sample preparation for SANS and XRD studies. Biodeuteration and SANS studies on Bio-SANS were supported by the OBER funded Center for Structural Molecular Biology (CSMB) under Contract FWP ERKP291, using the High Flux Isotope Reactor supported by the Office of Basic Energy Sciences, U. S. Department of Energy. SFG studies performed by S.H and S.H.K were supported by the Center for Lignocellulose Structure and Formation, an Energy Frontier Research Center, funded by DOE Office of Basic Energy Sciences, under Award DE-SC0001090. We thank Kevin Weiss for the technical assistance. The D_2O used in this research was supplied by United States Department of Energy Office of Science by the Isotope Program in the office of Nuclear Physics. This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this

manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan)

ASSOCIATED CONTENT Supplementary file 1

AUTHOR INFORMATION

Corresponding Author

*Corresponding authors: oneillhm@ornl.gov[‡], Present address: Neutron Scattering Division, Oak Ridge National Laboratory, TN:37830: davisonbh@ornl.gov Present address: BioSciences Division, Oak Ridge National Laboratory, TN:37830.

Present Addresses [l] Chemical Engineering Department, Aalto University, Finland FI-00076

All authors have given approval to the final version of the manuscript.

Abbreviations

hydromethyl furfural (5-HMF), Deuterium oxide (D_2O), XRD: X-ray diffraction; SFG: Sum frequency generation spectroscopy; SANS: Small angle neutron scattering; UV: Ultra Violet, Vis: Visible, QS: Quantitative saccharification; HPLC: high performance liquid chromatography; DAP: dilute acid pretreatment; R_g: radius of gyration; α: power law exponent

4.7). Supplementary information

4.7.1). Fits for the XRD profiles

Figure description (S4.7): The (100), (010), (110) and (012/102) peaks in the native cellulose, pretreated cellulose and pretreated xyloglucan-cellulose were fitted using pseudo-Voight function. For the native composites an extra Gaussian peak at 2θ = 21.5 was added.

4.7.2). Peak deconvolution of SFG spectra

4.7.3). Contrast variation series

4.7.4). Comparison of differently labelled bacterial cellulose

Figure description S4.9For the purpose of determining the contrast match point for glucomannan, three Q values were selected (-0.005, 0.007 and 0.01) along each curve which are shown by dashed black lines. The intensity values were noted at each Q point for all the curves and these were plotted against D2O concentration. This resulted in a I(Q) vs D2O concentration plot for each Q value which was fit to a parabola equation. The resulting minimum was calculated and that was the match point. As expected, three match point values were obtained and all those three were similar values

Figure Description S4.10: The scattering profile for hydrogenated cellulose in 100% D2O and partially deuterated cellulose in 45% D2O was measured to make sure that incorporation of deuterium in cellulose does not lead to a change in structure of cellulose. No significant change in the two profiles were observed.

Figure Description S4.11: As described in the materials and method section, the unified fit was used to fit all SANS curves. The q-range from 0.003 Å-1 to 0.3 Å-1 was used to fit the curve.

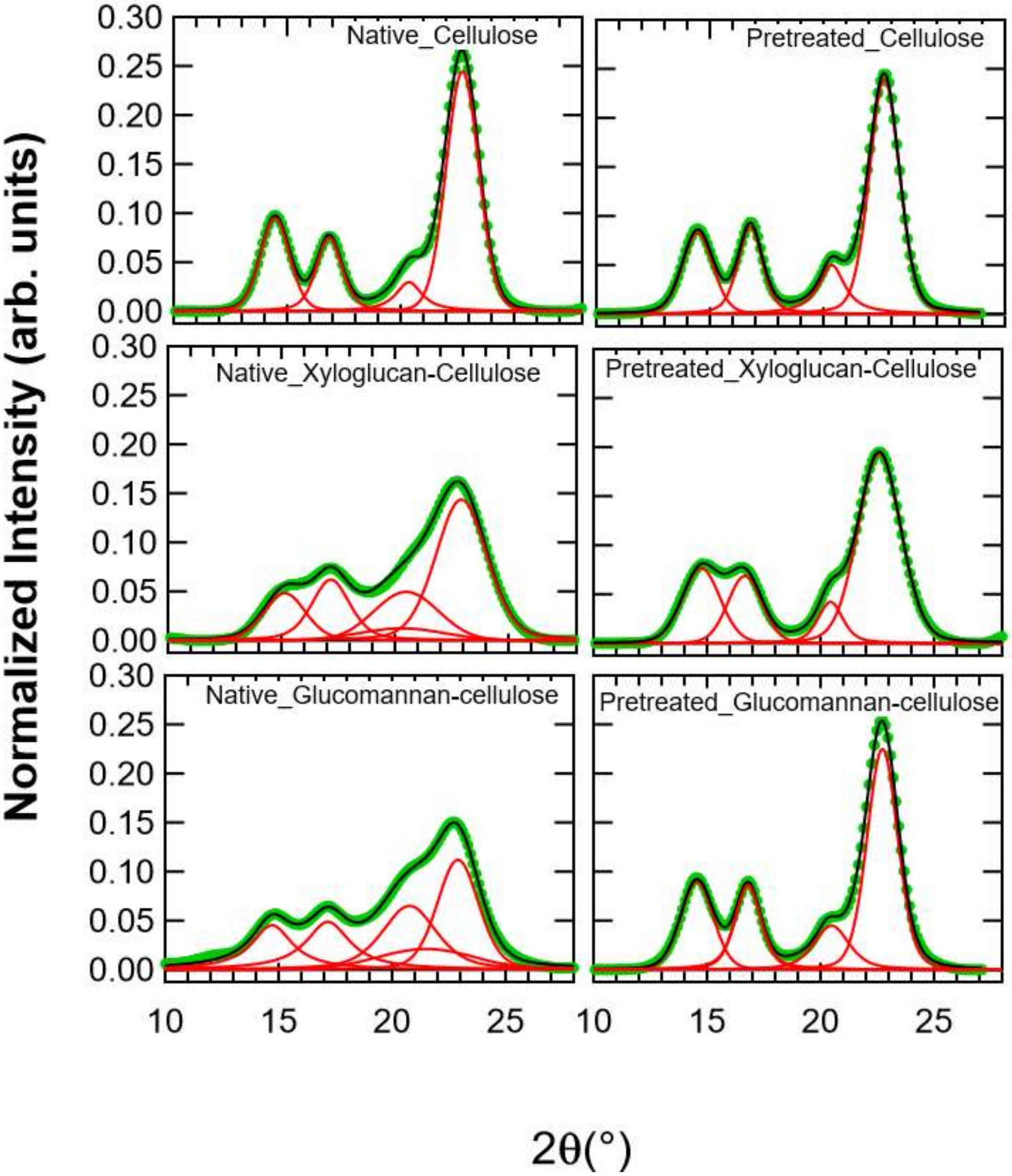


Figure 4.7: Wide angle x-ray scattering intensity profile (green curve) for native and pretreated cellulose and hemicellulose-cellulose composites with overall fit (black curve). Each profile was fit for background and 4 lattice peak -(100), (010), (110) and (012/102)

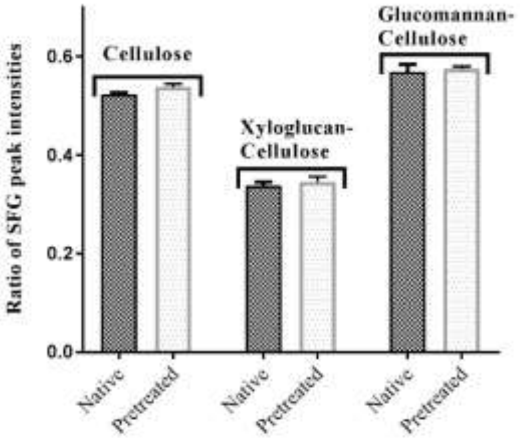


Figure 4.8: Plot showing Iα values determined by peak deconvolution of SFG spectra for native and pretreated cellulose and composites.

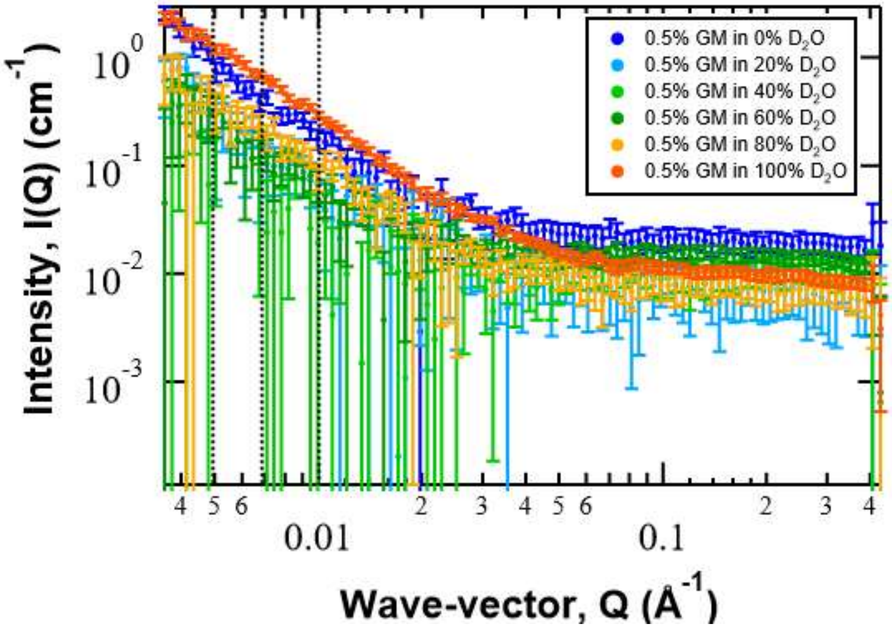


Figure 4.9: I(Q) vs Q plot for 0.5% glucomannan dissolved in solvent containing varying amount of D2O (0%, 20%, 40%, 60%, 80% and 100%).

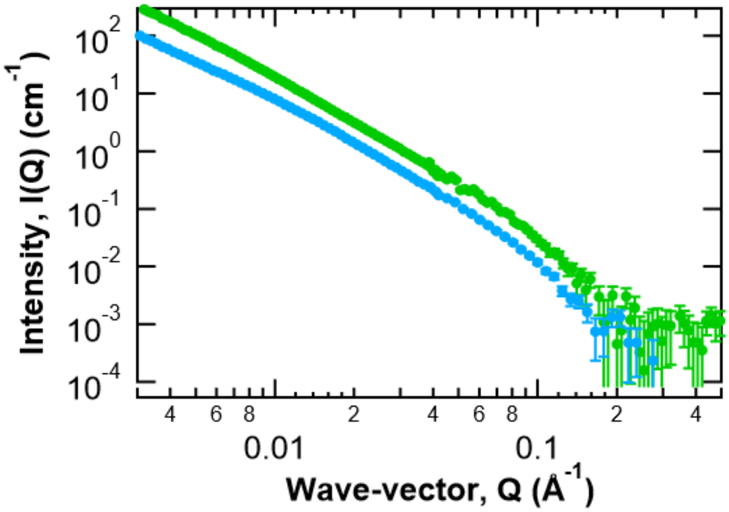


Figure 4.10: I(Q) vs Q plot for cellulose prepared in protiated media and measured in 100% D2O (green) and partially deuterated cellulose prepared in 100% D2O media measured with SANS in 45% D2O is shown for comparison.

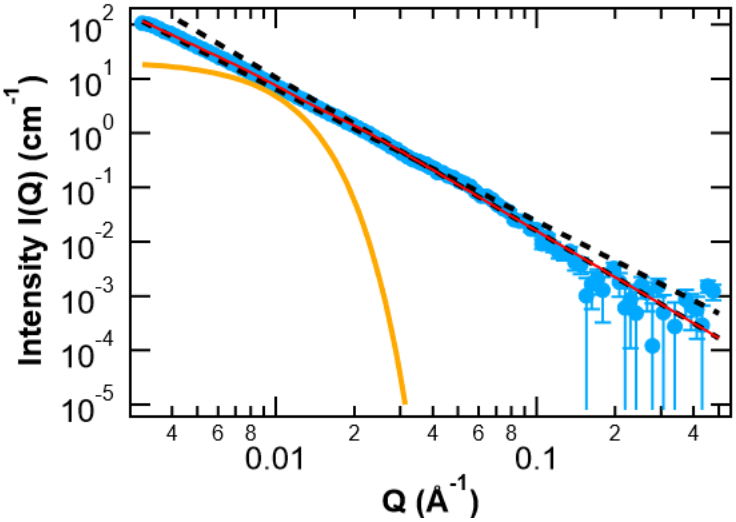


Figure 4.11: Overall fit (red) obtained for experimental intensity of native partially deuterated cellulose in 45% D2O (blue circle). Two levels were used and each term (Rg= orange; Power law = black dashed line) in eq 3 resulted in the overall fit

In order to obtain the overall fit, two levels with a Rg cut off for level 1 were used. Each level in turn comprised of two fitting functions: a Guinier function that gives information about particle size (Rg) and power law that gives information about the properties of the internal and/or external surface of the particle. The second level Rg and background were fixed to particular values during the fitting process. The parameter thus obtained were subjected to the uncertainty test which gives an estimate of error in parameter.

4.8). References

1. Naik, S. N.; Goud, V. V.; Rout, P. K.; Dalai, A. K., Production of first and second generation biofuels: a comprehensive review. *Renewable Sustainable Energy Rev.* **2010,** *14* (2), 578-597.
2. Perlack, R. D.; Eaton, L. M.; Turhollow Jr, A. F.; Langholtz, M. H.; Brandt, C. C.; Downing, M. E.; Graham, R. L.; Wright, L. L.; Kavkewitz, J. M.; Shamey, A. M., US billion-ton update: biomass supply for a bioenergy and bioproducts industry. **2011**.
3. Brodeur, G.; Yau, E.; Badal, K.; Collier, J.; Ramachandran, K. B.; Ramakrishnan, S., Chemical and Physicochemical Pretreatment of Lignocellulosic Biomass: A Review. *Enzyme Res.* **2011,** *2011*, 17.
4. Kumar, A. K.; Sharma, S., Recent updates on different methods of pretreatment of lignocellulosic feedstocks: a review. *Bioresources and bioprocessing* **2017,** *4* (1), 7.
5. Behera, S.; Arora, R.; Nandhagopal, N.; Kumar, S., Importance of chemical pretreatment for bioconversion of lignocellulosic biomass. *Renewable Sustainable Energy Rev.* **2014,** *36*, 91-106.
6. Talebnia, F.; Karakashev, D.; Angelidaki, I., Production of bioethanol from wheat straw: An overview on pretreatment, hydrolysis and fermentation. *Bioresour. Technol.* **2010,** *101* (13), 4744-4753.
7. Menon, V.; Rao, M., Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery concept. *Prog Energ Combust* **2012,** *38* (4), 522-550.
8. Esteghlalian, A.; Hashimoto, A. G.; Fenske, J. J.; Penner, M. H., Modeling and optimization of the dilute-sulfuric-acid pretreatment of corn stover, poplar and switchgrass. *Bioresour. Technol.* **1997,** *59* (2), 129-136.
9. Saha, B. C.; Iten, L. B.; Cotta, M. A.; Wu, Y. V., Dilute acid pretreatment, enzymatic saccharification and fermentation of wheat straw to ethanol. *Process Biochem* **2005,** *40* (12), 3693-3700.
10. Sassner, P.; Mårtensson, C.-G.; Galbe, M.; Zacchi, G., Steam pretreatment of H2SO4-impregnated Salix for the production of bioethanol. *Bioresour. Technol.* **2008,** *99* (1), 137-145.
11. Hsu, T.-C.; Guo, G.-L.; Chen, W.-H.; Hwang, W.-S., Effect of dilute acid pretreatment of rice straw on structural properties and enzymatic hydrolysis. *Bioresour. Technol.* **2010,** *101* (13), 4907-4913.
12. Pingali, S. V.; Urban, V. S.; Heller, W. T.; McGaughey, J.; O'Neill, H.; Foston, M.; Myles, D. A.; Ragauskas, A.; Evans, B. R., Breakdown of Cell Wall Nanostructure in Dilute Acid Pretreated Biomass. *Biomacromolecules* **2010,** *11* (9), 2329-2335.
13. Foston, M.; Ragauskas, A. J., Changes in lignocellulosic supramolecular and ultrastructure during dilute acid pretreatment of Populus and switchgrass. *Biomass Bioenergy* **2010,** *34* (12), 1885-1895.
14. Pu, Y.; Hu, F.; Huang, F.; Davison, B. H.; Ragauskas, A. J., Assessing the molecular structure basis for biomass recalcitrance during dilute acid and hydrothermal pretreatments. *Biotechnol. Biofuels* **2013,** *6* (1), 15.
15. Inouye, H.; Zhang, Y.; Yang, L.; Venugopalan, N.; Fischetti, R. F.; Gleber, S. C.; Vogt, S.; Fowle, W.; Makowski, B.; Tucker, M.; Ciesielski, P.; Donohoe, B.; Matthews, J.; Himmel, M. E.; Makowski, L., Multiscale deconstruction of molecular architecture in corn stover. *Sci. Rep.* **2014,** *4*, 3756.

16. Hu, F.; Jung, S.; Ragauskas, A., Pseudo-lignin formation and its impact on enzymatic hydrolysis. *Bioresour. Technol.* **2012,** *117*, 7-12.

17. Li, H.; Pu, Y.; Kumar, R.; Ragauskas, A. J.; Wyman, C. E., Investigation of lignin deposition on cellulose during hydrothermal pretreatment, its effect on cellulose hydrolysis, and underlying mechanisms. *Biotechnol Bioeng* **2014,** *111* (3), 485-492.

18. J., S. M.; Sridhar, V.; R., D. S.; P., T. M.; E., H. M.; B., V. T., Deposition of Lignin Droplets Produced During Dilute Acid Pretreatment of Maize Stems Retards Enzymatic Hydrolysis of Cellulose. *Biotechnol. Prog.* **2007,** *23* (6), 1333-1339.

19. Nishiyama, Y.; Kuga, S.; Okano, T., Mechanism of mercerization revealed by X-ray diffraction. *J Wood Sci* **2000,** *46* (6), 452-457.

20. Wada, M.; Chanzy, H.; Nishiyama, Y.; Langan, P., Cellulose IIII crystal structure and hydrogen bonding by synchrotron X-ray and neutron fiber diffraction. *Macromolecules* **2004,** *37* (23), 8548-8555.

21. Cerclier, C.; Cousin, F.; Bizot, H.; Moreau, C.; Cathala, B., Elaboration of Spin-Coated Cellulose-Xyloglucan Multilayered Thin Films. *Langmuir* **2010,** *26* (22), 17248-17255.

22. Benselfelt, T.; Cranston, E. D.; Ondaral, S.; Johansson, E.; Brumer, H.; Rutland, M. W.; Wagberg, L., Adsorption of Xyloglucan onto Cellulose Surfaces of Different Morphologies: An Entropy-Driven Process. *Biomacromolecules* **2016,** *17* (9), 2801-2811.

23. Falcoz-Vigne, L.; Ogawa, Y.; Molina-Boisseau, S.; Nishiyama, Y.; Meyer, V.; Petit-Conil, M.; Mazeau, K.; Heux, L., Quantification of a tightly adsorbed monolayer of xylan on cellulose surface. *Cellulose* **2017,** *24* (9), 3725-3739.

24. Martinez-Sanz, M.; Gidley, M. J.; Gilbert, E. P., Application of X-ray and neutron small angle scattering techniques to study the hierarchical structure of plant cell walls: a review. *Carbohydr Polym* **2015,** *125*, 120-34.

25. Martínez-Sanz, M.; Lopez-Sanchez, P.; Gidley, M. J.; Gilbert, E. P., Evidence for differential interaction mechanism of plant cell wall matrix polysaccharides in hierarchically-structured bacterial cellulose. *Cellulose* **2015,** *22* (3), 1541-1563.

26. Martinez-Sanz, M.; Gidley, M. J.; Gilbert, E. P., Hierarchical architecture of bacterial cellulose and composite plant cell wall polysaccharide hydrogels using small angle neutron scattering. *Soft Matter* **2016,** *12* (5), 1534-49.

27. Whitney, S. E. C.; Brigham, J. E.; Darke, A. H.; Reid, J. S. G.; Gidley, M. J., In-Vitro Assembly of Cellulose/Xyloglucan Networks - Ultrastructural and Molecular Aspects. *Plant J* **1995,** *8* (4), 491-504.

28. Martinez-Sanz, M.; Mikkelsen, D.; Flanagan, B. M.; Gidley, M. J.; Gilbert, E. P., Multi-scale characterisation of deuterated cellulose composite hydrogels reveals evidence for different interaction mechanisms with arabinoxylan, mixed-linkage glucan and xyloglucan. *Polymer* **2017,** *124*, 1-11.

29. Mikkelsen, D.; Gidley, M. J., Formation of Cellulose-Based Composites with Hemicelluloses and Pectins Using Gluconacetobacter Fermentation In *The Plant Cell Wall: Methods and Protocols*, Popper, Z. A., Ed. Humana Press: Totowa, NJ, 2011; Vol. 715, pp 197-208.

30. Lee, C. M.; Gu, J.; Kafle, K.; Catchmark, J.; Kim, S. H., Cellulose produced by Gluconacetobacter xylinus strains ATCC 53524 and ATCC 23768: Pellicle formation, post-synthesis aggregation and fiber density. *Carbohydr. Polym.* **2015,** *133*, 270-276.

31. Penttilä, P. A.; Imai, T.; Sugiyama, J., Fibrillar assembly of bacterial cellulose in the presence of wood-based hemicelluloses. *Int J Biol Macromol* **2017,** *102* (Supplement C), 111-118.

32. Brown, R. M., The biosynthesis of cellulose. *J Macromol Sci Pure* **1996,** *A33* (10), 1345-1373.

33. Brown, R. M.; Willison, J. H. M.; Richardson, C. L., Cellulose Biosynthesis in Acetobacter-Xylinum - Visualization of Site of Synthesis and Direct Measurement of Invivo Process. *P Natl Acad Sci USA* **1976,** *73* (12), 4565-4569.

34. Brown, R. M.; Saxena, I. M.; Kudlicka, K., Cellulose biosynthesis in higher plants. *Trends Plant Sci* **1996,** *1* (5), 149-156.

35. Tokoh, C.; Takabe, K.; Sugiyama, J.; Fujita, M., Cellulose synthesized by Acetobacter xylinum in the presence of plant cell wall polysaccharides. *Cellulose* **2002,** *9* (1), 65-74.

36. Whitney, S. E. C.; Wilson, E.; Webster, J.; Bacic, A.; Reid, J. S. G.; Gidley, M. J., Effects of structural variation in xyloglucan polymers on interactions with bacterial cellulose. *Am J Bot* **2006,** *93* (10), 1402-1414.

37. Park, Y. B.; Lee, C. M.; Kafle, K.; Park, S.; Cosgrove, D. J.; Kim, S. H., Effects of Plant Cell Wall Matrix Polysaccharides on Bacterial Cellulose Structure Studied with Vibrational Sum Frequency Generation Spectroscopy and X-ray Diffraction. *Biomacromolecules* **2014,** *15* (7), 2718-2724.

38. Iwata, T.; Indrarti, L.; Azuma, J. I., Affinity of hemicellulose for cellulose produced by Acetobacter xylinum. *Cellulose* **1998,** *5* (3), 215-228.

39. Uhlin, K. I.; Atalla, R. H.; Thompson, N. S., Influence of Hemicelluloses on the Aggregation Patterns of Bacterial Cellulose. *Cellulose* **1995,** *2* (2), 129-144.

40. Atalla, R. H.; Vanderhart, D. L., Native Cellulose - a Composite of 2 Distinct Crystalline Forms. *Science* **1984,** *223* (4633), 283-285.

41. Wada, M.; Okano, T.; Sugiyama, J., Allomorphs of native crystalline cellulose I evaluated by two equatorial d-spacings. *J Wood Sci* **2001,** *47* (2), 124-128.

42. O'Neill, H.; Shah, R.; Evans, B. R.; He, J. H.; Pingali, S. V.; Chundawat, S. P. S.; Jones, A. D.; Langan, P.; Davison, B. H.; Urban, V., Production of Bacterial Cellulose with Controlled Deuterium-Hydrogen Substitution for Neutron Scattering Studies. *Method Enzymol* **2015,** *565,* 123-146.

43. Ren, J. L.; Sun, R. C.; Liu, C. F.; Cao, Z. N.; Luo, W., Acetylation of wheat straw hemicelluloses in ionic liquid using iodine as a catalyst. *Carbohydr. Polym.* **2007,** *70* (4), 406-414.

44. He, J.; Pingali, S. V.; Chundawat, S. P.; Pack, A.; Jones, A. D.; Langan, P.; Davison, B. H.; Urban, V.; Evans, B.; O'Neill, H., Controlled incorporation of deuterium into bacterial cellulose. *Cellulose* **2014,** *21* (2), 927-936.

45. A. Sluiter, B. H., R. Ruiz, C. Scarlata, J. Sluiter, and D. Templeton Determination of Sugars, Byproducts, and Degradation Products in Liquid Fraction Process Samples. *Technical Report NREL/TP-510-42623* **2006**, 1-11.

46. Ilavsky, J.; Jemian, P. R., Irena: tool suite for modeling and analysis of small-angle scattering. *J Appl Crystallogr* **2009,** *42,* 347-353.

47. Bruckner, S., Estimation of the background in powder diffraction patterns through a robust smoothing procedure. *J Appl Crystallogr* **2000,** *33* (2), 977-979.

48. Wojdyr, M., Fityk: a general-purpose peak fitting program. *J Appl Crystallogr* **2010,** *43,* 1126-1128.

49. Garvey, C. J.; Parker, I. H.; Simon, G. P., On the interpretation of X-ray diffraction powder patterns in terms of the nanostructure of cellulose I fibres. *Macromol Chem Physic* **2005,** *206* (15), 1568-1575.

50. Lee, C. M.; Kafle, K.; Park, Y. B.; Kim, S. H., Probing crystal structure and mesoscale assembly of cellulose microfibrils in plant cell walls, tunicate tests, and bacterial films using vibrational sum frequency generation (SFG) spectroscopy. *Phys. Chem. Chem. Phys.* **2014,** *16* (22), 10844-10853.
51. Heller, W. T.; Cuneo, M.; Debeer-Schmitt, L.; Do, C.; He, L.; Heroux, L.; Littrell, K.; Pingali, S. V.; Qian, S.; Stanley, C.; Urban, V. S.; Wu, B.; Bras, W., The suite of small-angle neutron scattering instruments at Oak Ridge National LaboratoryThis article will form part of a virtual special issue on advanced neutron scattering instrumentation, marking the 50th anniversary of the journal. *J Appl Crystallogr* **2018,** *51* (2), 242-248.
52. Berry, K. D.; Bailey, K. M.; Beal, J.; Diawara, Y.; Funk, L.; Hicks, J. S.; Jones, A. B.; Littrell, K. C.; Pingali, S. V.; Summers, P. R.; Urban, V. S.; Vandergriff, D. H.; Johnson, N. H.; Bradley, B. J., Characterization of the neutron detector upgrade to the GP-SANS and Bio-SANS instruments at HFIR. *Nucl Instrum Meth A* **2012,** *693*, 179-185.
53. Beaucage, G., Small-angle scattering from polymeric mass fractals of arbitrary mass-fractal dimension. *J Appl Crystallogr* **1996,** *29* (2), 134-146.
54. Park, S.; Baker, J. O.; Himmel, M. E.; Parilla, P. A.; Johnson, D. K., Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnol. Biofuels* **2010,** *3* (1), 10.
55. Gu, J.; Catchmark, J. M., The impact of cellulose structure on binding interactions with hemicellulose and pectin. *Cellulose* **2013,** *20* (4), 1613-1627.
56. Kent, M. S.; Cheng, G.; Murton, J. K.; Carles, E. L.; Dibble, D. C.; Zendejas, F.; Rodriquez, M. A.; Tran, H.; Holmes, B.; Simmons, B. A.; Knierim, B.; Auer, M.; Banuelos, J. L.; Urquidi, J.; Hjelm, R. P., Study of Enzymatic Digestion of Cellulose by Small Angle Neutron Scattering. *Biomacromolecules* **2010,** *11* (2), 357-368.
57. Langan, P.; Petridis, L.; O'Neill, H. M.; Pingali, S. V.; Foston, M.; Nishiyama, Y.; Schulz, R.; Lindner, B.; Hanson, B. L.; Harton, S.; Heller, W. T.; Urban, V.; Evans, B. R.; Gnanakaran, S.; Ragauskas, A. J.; Smith, J. C.; Davison, B. H., Common processes drive the thermochemical pretreatment of lignocellulosic biomass. *Green Chem* **2014,** *16* (1), 63-68.
58. DeMartini, J. D.; Pattathil, S.; Avci, U.; Szekalski, K.; Mazumder, K.; Hahn, M. G.; Wyman, C. E., Application of monoclonal antibodies to investigate plant cell wall deconstruction for biofuels production. *Energy Environ. Sci.* **2011,** *4* (10), 4332-4339.
59. O'Neill, H.; Pingali, S. V.; Petridis, L.; He, J.; Mamontov, E.; Hong, L.; Urban, V.; Evans, B.; Langan, P.; Smith, J. C.; Davison, B. H., Dynamics of water bound to crystalline cellulose. *Sci. Rep.* **2017,** *7* (1), 11840.
60. Park, Y. B.; Cosgrove, D. J., Xyloglucan and its Interactions with Other Components of the Growing Cell Wall. *Plant Cell Physiol* **2015,** *56* (2), 180-194.
61. Zhou, Q.; Rutland, M. W.; Teeri, T. T.; Brumer, H., Xyloglucan in cellulose modification. *Cellulose* **2007,** *14* (6), 625-641.

Reference for Supplement:

1. A. Sluiter, B. H., R. Ruiz, C. Scarlata, J. Sluiter, and D. Templeton Determination of Sugars, Byproducts, and Degradation Products in Liquid Fraction Process Samples. Technical Report NREL/TP-510-42623 2006, 1-11.

2. Dumitrache, A.; Natzke, J.; Rodriguez, M.; Yee, K. L.; Thompson, O. A.; Poovaiah, C. R.; Shen, H.; Mazarei, M.; Baxter, H. L.; Fu, C. X.; Wang, Z. Y.; Biswal, A. K.; Li, G. F.; Srivastava, A. C.; Tang, Y. H.; Stewart, C. N.; Dixon, R. A.; Nelson, R. S.; Mohnen, D.; Mielenz, J.; Brown, S. D.; Davison, B. H., Transgenic switchgrass (Panicum virgatum L.) targeted for reduced recalcitrance to bioconversion: a 2-year comparative analysis of field-grown lines modified for target gene or genetic element expression. Plant Biotechnol J 2017, 15 (6), 688-697.

Chapter 5). Structural insights into low and high recalcitrance natural poplar variants using neutron and x-ray scattering

A version of this chapter will be published in a peer-reviewed journal as follows:

Shah R., Bhagia S., Keum J, Pingali S.V., Ragauskas A.J, Davison B.H and O’Neill H “Structural insights into low and high recalcitrance natural poplar variants using neutron and x-ray scattering”, ACS Sustainable Chemistry and Engineering (to be submitted)

5.1). Abstract

While a correlation between physiochemical properties of lignin and amount of sugar release is often observed, the changes in cell wall structure due to differing lignin content and composition have not been studied. We used small-angle neutron scattering (SANS) and wide-angle x-ray scattering (WAXS) to investigate the structural changes in the cell wall of two naturally occurring variants of poplar species *Populus trichocarpa*, BESC-316 and GW-11012 with klason lignin content 17.8 and 23.2% respectively, both before and after hot water pretreatment. SANS results show aggregated cellulose microfibril arrangement in GW-11012 while well-ordered cellulose microfibrils in BESC-316. Lignin was distributed in the cell wall of BESC-316 while in GW-11012 it was aggregated to size of 52 Å. Post-pretreatment, the cellulose was coalesced and formed similar size fibrils while lignin aggregated in both with slight difference in lignin aggregate size at 180 ˚C/45min. The crystallite size and orientation were both larger and better arranged in GW-11012 than BESC-316 and this trend was maintained post pretreatment. Our study shows that deposition of polymers in the cell wall is co-dependent, as structural changes in lignin distribution and in cellulose crystallite size and orientation were simultaneously observed. Such changes may contribute in cellulose surface accessible for enzymatic hydrolysis.

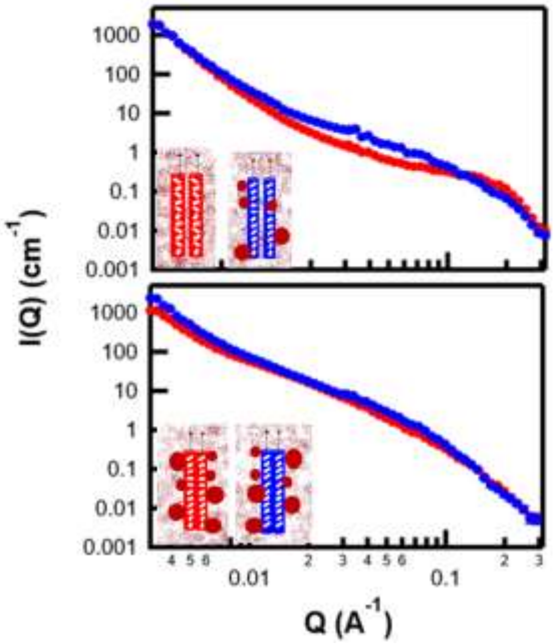


Figure 5.1: Differences in cellulose microfibrils and lignin distribution between BESC-316 (red) and GW-11012 (blue) in native (top) and after 180 ˚C/45min hot water pretreatment (bottom)

5.2). Introduction

Lignocellulosic biomass is primarily composed of the carbon-rich polymers cellulose, hemicellulose, and lignin, making it an excellent source of carbon-based fuels and bioproducts.[1-2] Thermochemical pretreatment approaches are commonly used as an initial step to disrupt the rigid cell wall structure to enable greater access to the hydrolytic enzymes for conversion cellulose to glucose that is then fermented to fuels such as ethanol or butanol. However, even after the initial disruption of the cell wall structure by pretreatment, enzymatic digestion is not as efficient as expected, rendering the commercialization of sustainable biofuels costlier as compared to petroleum-based fuels.[3] Several factors are thought to play a role in biomass recalcitrance that include cell wall porosity,[4] lignin composition and linkages,[5] cellulose crystallinity,[6] and also hemicellulose content and structure.[7-8] Of these factors, lignin is often cited as a major reason for biomass recalcitrance in both native and thermochemically pretreated biomass.[9-12] It has been shown to block enzyme access to the cellulose surface,[13] and in pretreated solids it unproductively bind to enzymes decreasing their effectiveness in biomass digestion .[6, 14] Also, transgenic species that have been genetically modified to reduce lignin content have been reported to be more susceptible to enzymatic hydrolysis than their naturally occurring counterparts, supporting that lignin negatively effects sugar release.[15]

In addition to the amount of lignin, that can vary significantly among woody plants (10-30%) [16], lignin composition is also an important determinant in biomass recalcitrance. It is typically synthesized in secondary cell walls and is composed of varying ratios of guaiacyl (G), syringyl (S), *p*-hydroxymethyl (H) hydroxycinnamyl monomers. Variability in lignin composition and content has been examined to find a correlation between lignin content and composition and sugar release.[17] A study by Studer et al. characterized 47 *Populus trichocarpa* phenotypes that differed in lignin content and S/G ratio that were part of an undomesticated *P. trichocarpa* population with 1100 individual variants.[18] Poplars from this subset were tested for total sugar release using enzymatic digestion alone and using a combined liquid hot water pretreatment and enzymatic digestion. A negative correlation for glucose release was found for pretreated samples with an S/G ratio less than 2 whereas phenotypes with a higher S/G ratio displayed greater glucose release. In addition, for enzymatic digestion without pretreatment, sugar release increased significantly when lignin content was <20% irrespective of lignin S/G ratio. However, the authors also noted that there were several outlier phenotypes with exceptional sugar release that featured average lignin content and S/G ratio suggesting that lignin content and composition are not the only factor to influence sugar release.[18] Another study investigated poplar variants that carried a rare natural mutation in the *P. trichocarpa* 5-Enolpyruvylshikimate 3-Phosphate Synthase gene that is linked to reduced lignin biosynthesis and affected recalcitrance of the field-grown plants.[19-20] Sugar yields were analyzed after different processing conditions and glucan yields were significantly higher from the rare poplar variants compared to a high lignin comparator. However, relative glucan yields from the individual variants did not show a clear trend with pretreatment severity suggesting that other factors may influence sugar release, similar to the conclusions of Studer et al. as described above.[18] These studies highlight the critical need for a better understanding of the cell wall structure in order to deconvolute the processes that occur during cell wall deconstruction.

Insights into the molecular organization and architecture of the plant cell wall have been obtained using microscopy, spectroscopy and imaging techniques.[21] Combining information from complimentary techniques is vital for a complete understanding of cell wall ultrastructure. For instance, atomic force microscopy, a type of scanning probe microscopy, has been used for observing nanoscale surface features of lignocellulosic biomass such as cell wall morphology and surface roughness.[22] Coherent anti-stokes Raman scattering and electron microscopy was used to image differences in lignin distribution for cell wall sections from wildtype alfalfa and a variant with downregulated lignin synthesis,[23] and also tension wood from poplar.[24] Both the lignin-downregulated alfalfa variant and tension wood showed the presence of lignin aggregates in contrast to the wildtype alfalfa, in which lignin was evenly distributed in the cell wall. Solid state nuclear magnetic resonance spectroscopy provides useful insights into cell wall polymer interactions and showed that xylan bound to cellulose has a different conformation than in solution[25], and lignin makes electrostatic interactions with xylan.[26] X-ray and neutron scattering techniques have been also used to study the plant cell wall.[27] Small angle scattering in particular is a powerful tool to study the intact cell walls with minimal sample preparation to obtain structural information in the length scale spanning from one to several hundred nanometers. Structural information such as distance between cellulose microfibrils, cross-sectional size of cellulose microfibrils, distribution of the matrix co-polymers, and overall cell wall morphology can be determined using these techniques.[27]

Here, we report on the cell wall structure of BESC-316 and GW-11012, two natural poplar variants from the poplar genome-wide association study.[19] These two variants were chosen for analysis because a previous study[20] found that native GW-11012 released significantly more glucose per cellulose content compared to BESC-316 after enzymatic hydrolysis. This trend continued after liquid hot water (LHW) pretreatment of the variants at different pretreatment temperatures and severities. After LHW pretreatments performed at a severity factor (log R_0) of 3.6 (160 ˚C for 68 min and 180 ˚C for 18min) GW-11012 released approximately 1.3 times more sugar compared to BESC-316 (Table S1). At a higher severity factor (R_0 = 4, 180 ˚C/ 45min) the amount of sugar released was even greater for GW-11012, 1.6 times, compared to BESC-316 (Table S1). A major difference between the two variants was in their lignin content and composition. BESC-316, the high lignin comparator had a lignin content of 23.2% and S:G ratio of 1.67. On the other hand, GW-11012 is a low lignin poplar variant with a lignin content of 17.8% with S:G ratio of 3.0.[28]

It was of interest to determine if there was a structural basis for the differences in recalcitrance between the two poplar variants. We used a combination of small-angle neutron scattering (SANS) to characterize the nanoscale (~1 – 100 nm) properties of the wood and X-ray diffraction (XRD) to obtain atomic level structural information about cellulose structure. The structural changes that occurred in the biomass samples during pretreatment were recorded in real-time using a specially designed pressure cell for time-resolved SANS pretreatment studies. Structural characterization was complemented by Simons' stain to determine cellulose accessibility~~.~~ Significant differences in both cellulose organization and lignin distribution were observed in native BESC-316 and GW-11012. Although time-resolved SANS studies showed differences in the progression of lignin aggregate formation and cellulose coalescence during pretreatment at different severities for both poplar variants, it did not fully explain the increased digestibility of GW-11012. On the other hand, our data support that differences observed at the atomic scale between the cellulose organization of the two variants may be important in reducing the recalcitrance of the biomass. Our results provide new insight into the interplay of cell wall polymer interactions in both cell wall synthesis and deconstruction and may offer new directions for overcoming biomass recalcitrance.

5.3). Materials and methods:

5.3.1). Sample preparation for small angle neutron scattering and X-ray diffraction

Four-year-old *Populus trichocarpa* BESC-316 and GW-11012 obtained from a field site in Clatskanie, Oregon were used for this study.[55] The estimated moisture content was determined by IR moisture analyzer and was found to be 8.82% and 6.92% for BESC-316 and GW-11012, respectively. The outermost growth ring of both wood stems were cut longitudinal to the growth direction for X-ray diffraction and SANS studies. The cut wood pieces were 14 x 14 mm in size with a thickness of approx. 1mm. The wood pieces were incubated in D_2O or 35% D_2O solvent for at least 24 hr which included at least three exchanges into the fresh solvent to ensure complete D/H exchange. Delignification of native and pretreated wood samples was carried out as previously described. [56]

5.3.2). Small-angle neutron scattering data collection and analysis

SANS measurements on native and pretreated BESC-316 and GW-11012 were performed at the Bio-SANS instrument located at the High Flux Isotope Reactor facility in Oak Ridge National Laboratory (ORNL).[40] The main detector array was positioned at 15.5 m from the sample position, and the wing detector array was positioned 1.13 m from the sample position at an angle of 1.4° from the direct beam. Using this configuration, the Q range spanning $0.003 < Q\ (\text{Å}^{-1}) < 0.8$ was obtained using 6 Å neutrons with a relative wavelength spread ($\Delta\lambda/\lambda$) of 15%.

For real-time SANS pretreatment measurements, the samples were equilibrated in D_2O solvent and placed in a pressure cell with a 1 mm path length. The pressure cell used for these measurements is the enhanced angle pressure (EAP) cell, a modification of the McHugh Cell that allows measurement of scattering wave-vectors up to 0.6 Å^{-1}.[39] Each wood chip was aligned along its growth direction perpendicular to the neutron beam. The scattering intensity profile, I(Q) versus Q, for each sample was obtained by azimuthally averaging the processed 2D images which were normalized to incident beam monitor counts, corrected for detector dark current, pixel sensitivity and scattering from backgrounds such as solvent and sapphire windows of the pressure cell. LHW pretreatment reactions were performed for BESC-316 and GW-11012 at pretreatment severity factor[18] log R_0 = 3.6 for which the time and temperature conditions were 160 ˚C for 68 min (160˚C/68min) and 180 ˚C for 18 (180˚C/18 min), and at a higher severity factor (log R_0 = 4.0) at 180˚C for 45min (180˚C/45min) These conditions were selected based on differences in sugar release between the two variants which were previously reported by Bhagia et al.[18] After sealing the samples that were equilibrated in 100% D_2O solvent in EAP cell they were heated to the retention temperature of 160 ˚C or 180 ˚C within 4 min or 6 mins respectively, and after the retention time was completed they were then cooled to 25 ˚C within ~20 mins (see Figure 5, panels A-C).

The SANS data of delignified native and pretreated samples were collected at the EQ-SANS beamline[57] located at the Spallation Neutron Source (SNS) which was operating at 60 Hz. A wood chip was packed into a 1 mm pathlength titanium cell and filled with 100% D_2O. The neutron beam was collimated using an aperture size of 10 mm and the scattering data was collected by using sample of detector distance (SDD) of 2.5 m and 4 m. The neutron wavelength used for the two SDDs was 2.5 and 10 Å respectively. The combination of the SDD and the range of neutron wavelengths enabled us to obtain a Q range spanning $0.005 \text{ Å}^{-1} < Q < 0.7 \text{ Å}^{-1}$. The data was reduced by using the standard procedures implemented in the Mantid software[58] and an output file containing Q, I(Q) and error in I(Q) and Q was obtained. The intensity I(Q) in the output file was calibrated to be on absolute scale using porous silica as a standard.

The SANS data were fit to the multi-level Unified Fit model[31] implemented in the IRENA package[36] of Igor Pro software by Wavemetrics. As shown in Equation 1 the SANS intensity profile is a summation of individual levels (i) with each level modeled as the sum of an exponential and a power-law behavior

$$I(q) = \sum_{i} I_i(q) = G_i e^{-\frac{q^2 R_{g_i}^2}{3}} + e^{-\frac{q^2 R_{gCO_{i-1}}^2}{3}} \times B_i \left\{ \frac{\left(\mathrm{erf}\left(\frac{qR_{g_i}}{\sqrt{6}}\right)\right)^3}{q} \right\}^{P_i} + C \tag{5.1}$$

where R_{g_i} is the radius of gyration of the particle of the *i*th structural level, P_i is the power-law exponent of the *i*th structural level, G_i is the scalar for the Guinier function of the *i*th structural level, and B_i is the scalar for the power-law function of the *i*th structural level. Additionally, $R_{gCO_{i-1}}$ is the cut-off length scale of the power-law behavior of the *i*th structural level and *C* is the Q-independent constant background intensity. For BESC-316, the peak position (ζ) and the packing density parameter (k) were obtained from the Unified Fit analysis, as described previously.[38, 59]

5.3.3). X-ray diffraction data collection and analysis

XRD measurements were carried out using an Anton-Paar SAXSess mc^2 instrument with point collimated monochromatic CuK$_\alpha$ radiation ($\lambda = 1.54$ Å) operating at 50 kA and 40 kV. The beam size was 0.8 mm x 0.8 mm and the sample to detector distance was 42.1 mm. The wood pieces were aligned perpendicular to the direction of growth in the X-ray beam and the 2-dimensional (2D) scattering intensity image was collected using a multi-sensitive phosphorous screen with a pixel size of 0.042 mm. A cyclone plus storage phosphorous system (PerkinElmer) was used to read the intensity of the image plate. The background scattering of each image was determined by collecting image with no sample in the beam path and this 2D image was subtracted from the sample scattering image. The data collection time for each sample was 24 hr. The diffraction pattern showed the typical anisotropic scattering of the wood.

The azimuthal intensity distribution of the reflection (200) was calculated and used in Equation 2 and 3 to obtain the Hermann's orientation factor.[42-43]

$$\langle cos^2 \emptyset_{200} \rangle = \frac{\int_0^{90} I_{200}(\emptyset)\, cos^2 \emptyset \sin \emptyset \ d\emptyset}{\int_0^{90} I_{120}(\emptyset) \sin \emptyset \ d\emptyset} \tag{5.2}$$

$$f_{200} = \frac{3\,\langle cos^2 \emptyset_{200} \rangle - 1}{2} \tag{5.3}$$

As expected, the reflection planes (1-10), (110) and (200) were observed on the equatorial axis because the samples were aligned perpendicular to the growth direction of the samples.[60] The peak position for (200) was used to generate the intensity distribution along the azimuthal angle (Φ) from the background subtracted 2D pattern. The I(Φ) vs Φ profile was used with the Hermann's orientation function to determine the Hermann's orientation factor.

1D profile of I(Θ) vs 2Θ were made using the equatorial reflections. Peak fitting was performed to determine the crystallite size and crystallinity. The crystallinity was determined using Equation 4.

$$\%\, Crystallinity = \frac{(total\ scattering\ area - amorphous\ area)}{total\ scattering\ area} \times 100 \tag{5.4}$$

where 'total scattering area' and 'amorphous area' are the integrated areas under the curves in the 2θ range from 10˚ to 60˚. The crystallite size along the (200) was determined using the Scherrer equation shown in Equation 5

$$t = (0.9 \times \lambda) \div (\beta \times \cos\theta) \tag{5.5}$$

where, t is the size of the crystallite, λ is the wavelength of the incident X-rays, β and 2θ are the values for FWHM and peak position, respectively.

5.4). Results

5.4.1). Nanoscale structural changes for native and pretreated low and high lignin poplar cell wall:

5.4.1.1). Structural variations in native BESC-316 and GW-11012

Small angle neutron scattering was used to study the morphological changes in the cell wall structure of native and liquid hot water pretreated BESC-316 and GW-11012. The 2-dimensional (2D) SANS detector images of the scattering profiles of intact pieces of BESC-316 and GW-11012 in 100% D_2O solvent, after corrections and background subtraction, are shown in Figure 5.2.

At this solvent condition, maximum contrast is attained between the lignocellulose matrix and the solvent. The anisotropic scattering pattern is due to the alignment of the wood pieces perpendicular to the direction of growth in the neutron beam. The differences in the cell wall structures of the two variants are evident from the 2D images. BESC-316 shows a lobe-like feature that indicates the presence of correlated scattering particles [30] while GW-11012 has a sharp streak-like scattering pattern indicating well aligned scattering particles with non-uniform interparticle distance [31] (Figure 5.2). These differences between the samples in the scattering pattern can be related to differences in the arrangement of the cellulose microfibrils and matrix copolymers.

The 1-dimensional (1D) SANS profiles of BESC-316 and GW-11012 in 100% D_2O, are shown in Figure 5.3. The Unified Fit approach,[32] previously applied for studying intact biomass[33-34] and purified biomass polymers,[35] was used to fit the SANS data.

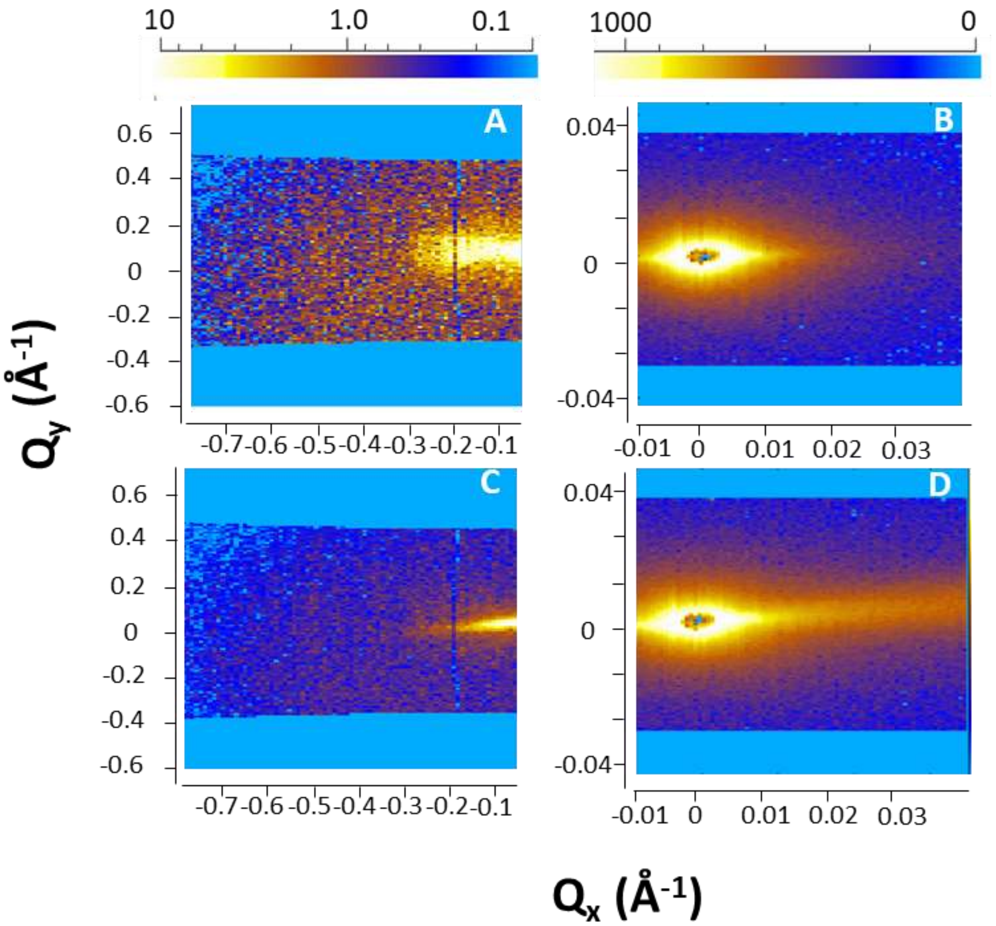


Figure 5.2: Two-dimensional detector intensity images are plotted as a function of Q_x vs. Q_y obtained from the dual detector configuration of the Bio-SANS instrument.[28]Panels A and B are from BESC-316 and panels C and D are from GW-11012. Panels A and C cover the Q range of 0.03 to 0.6 $Å^{-1}$ and panels B and D cover the Q range of 0.003 to 0.03 $Å^{-1}$. The intensity scale bar is on a logarithmic scale. Samples were equilibrated in 100% D_2O before measurement

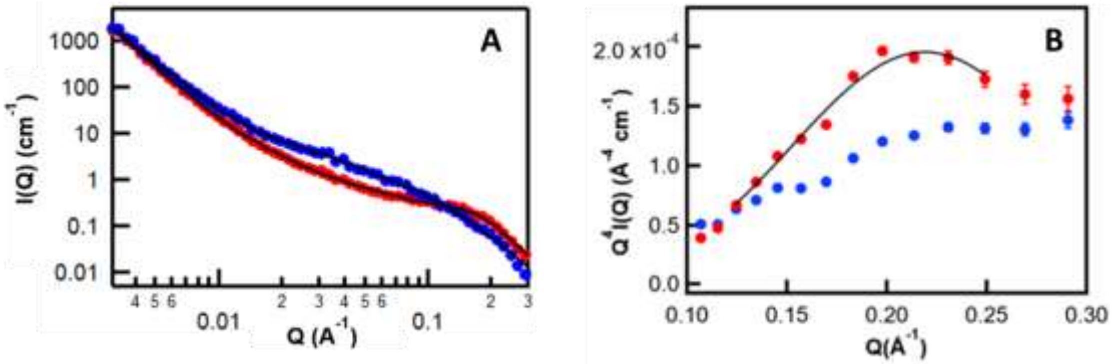


Figure 5.3: A). SANS scattering profile of native BESC-316 (red curve) and GW-11012 (blue curve) in 100% D_2O. The black solid line is the fit to the scattering curve obtained by using the Unified fitting model B). $Q^4 \cdot I(Q)$ vs. Q representation of the data to highlight distinguishable features in the high-Q region. The curve for BESC-316 was fit with a Gaussian fit (black)

Three structural levels were obtained from the fitting which covered the Q range from 0.1 Å^{-1} < Q < 0.4 Å^{-1}, 0.01 Å^{-1} < Q < 0.1 Å^{-1} and 0.003 < Q < 0.01 Å^{-1}.The fit values for the 1st structural level, that is dominated by the cellulose contribution to the SANS profiles, are summarized in Table 5.1. For BESC-316, radius of gyration (R_g) and a correlation function were used for curve fitting. The level-1 R_g value, 11 ± 2 Å, is interpreted as the cross-sectional dimension of a cellulose microfibril and is similar in size to the values obtained for cellulose microfibrils in previous studies. [31, 36-37] The peak position (ζ), 31 ± 5 Å, obtained from the correlation function, is the center to center distance of two adjacent microfibrils. The cross-sectional diameter of the cellulose microfibril, calculated to be 25 Å ($D = \sqrt{\frac{16}{3}}\, R_g$) indicates that the adjacent microfibrils are closely packed in the cell wall. The packing density parameter (k) quantifies the degree of packing of correlated particles. The possible values range from zero, for randomly packed particles, to 5.92 for a perfectly packed system such as a hexagonal or face centered cubic system.[32, 38] Here, the k value obtained was 1.3 ± 0.9, which is indicative of weak packing of cellulose microfibrils with interactions that do not extend past the next-neighbor correlations. In contrast, for the GW-11012 SANS curve, a single R_g was sufficient to get a good fit in the Level-1 structural level. The R_g value (21 ± 4 Å, D = 48 ± 9 Å) is almost double the value obtained for BESC-316, indicating that adjacent cellulose microfibrils are aggregated in GW-11012. And importantly, the absence of a peak at 0.2 Å^{-1} is consistent with not including a correlation function in the fit to level-1 structural regime of GW-11012 (Figure 5.3B).

The mid Q region (0.01 Å^{-1} > Q > 0.1 Å^{-1}) of the SANS curves provides information about the arrangement of the matrix co-polymers in the cell wall[33, 39] and similar to the high Q region, there were significant differences observed between native BESC-316 and GW-11012.

Table 5.1: Fitting parameters obtained for the SANS curves of native BESC-316 and GW-11012 using the Unified Fit approach

Sample	Level 1			Level2		Level3
	R_g (Å)	ETA (ζ) (Å)	Pack (k)	$2R_g$ (Å)	2P	
BESC-316	11 +/- 2	31 +/- 5	1.3 +/- 0.9	-	1.8 +/- 0.2	4.2 +/- 0.1
GW-11012	21 +/- 4	-	-	56 +/- 10	-	3.6 +/- 0.1

Error bars were obtained by performing uncertainty analysis in the Unified Fit [32, 38]

For BESC-316, this region of the SANS curve was fit with a single power-law that yielded an exponent of 1.8 ± 0.2. This value is similar to those obtained in previous studies for analysis of SANS data from poplar wood and is interpreted as the network of matrix copolymers in the cell wall. [31, 39] In contrast, the scattering intensity in the mid-Q region for GW-11012 exhibits a shoulder feature and curve fitting yielded a R_g of 56 ± 10 Å. The SANS profiles for GW-11012 before and after delignification are shown in Figure 5.4A. The scattering intensity is decreased in the mid-Q region providing evidence that this feature is related to lignin and shows a significant difference in the distribution of lignin between the two variants (Figure 5.4A).

Previous studies have also reported a similar feature in SANS profiles of pretreated biomass that was attributed to lignin aggregates in the cell wall.[31, 33] SANS analysis of BESC-316 and GW-11012 were also performed for samples equilibrated in 35% D_2O, in which cellulose scattering is suppressed.[34] The BESC-316 scattering curve shows a monotonic increase in scattering intensity over the measured Q-range and was fit to a combination of two power-laws as detailed in Table 5.2. Interestingly, a size feature is evident in GW-11012 at ~0.04 $Å^{-1}$ that is not present in the BESC-316 curve; the fit yielded a R_g of 46 ± 3 Å. A similar sized feature was observed in the 100% D_2O measurement of GW-11012 sample suggesting that this feature highlights the difference in the distribution of lignin between the two variants.

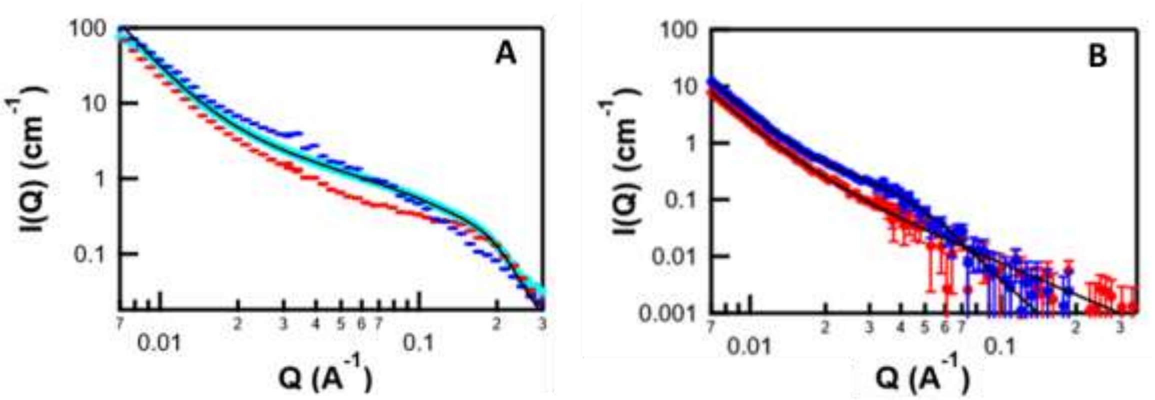


Figure 5.4: A). Comparison of SANS data of GW-11012 before (blue dots) and after delignification (cyan curve) in 100% D_2O solvent. BESC-316 (red dots) in 100% D_2O solvent is included for comparison. B). Comparison of SANS profiles of native BESC-316 (red) and GW-11012 (blue) in 35% D_2O solvent. In both panels, the solid black line is the fit obtained to the curves with the Unified fitting model.

Table 5.2: Fitting parameters obtained for the SANS profiles of native BESC-316 and GW-11012 in 35% D_2O using the Unified Fit approach

	Level 1		Level 2
	R_g (Å)	P	P
BESC-316	1120*	1.9 +/- 0.21	4.0 +/- 0.2
GW-11012	46 +/- 3	4*	3.7 +/- 0.1

The range of fitting for level 1 was 0.015 Å^{-1} < Q < 0.2 Å^{-1} and for level 2 was 0.001 Å^{-1} < Q <0.015 Å^{-1}

*the value for the parameter was fixed during the fitting process

The surface morphology of the cell wall is reflected by the power-law exponent (α) of the third structural level (Figure 5.3, Table 5.1). The value of $\alpha = 3$ is interpreted as the surface fractal with rough interfaces with the solvent. As the value increases and reaches 4.2 ± 0.1, the surface fractal is considered smooth. Such smooth surface is seen for BESC-316. The 3rd level α -value for GW-11012 is 3.6 ± 0.1 indicating a rough surface. Surface roughness has previously been associated with increased accessibility to enzymes.[31]

5.4.1.2). Real-time structural analysis of BESC-316 and GW-11012 during pretreatment

The structural changes that occur in biomass during pretreatment reactions were monitored *in-situ* by SANS using a pressure cell that mimics thermochemical pretreatment conditions typically performed in Parr reaction vessels. [40-41] Liquid hot water (LHW) pretreatment reactions were performed for BESC-316 and GW-11012 at retention temperatures of 160 ˚C for 68 min (160˚C/68min), and 180 ˚C for 18 and 45 min (180˚C/18 min, 180˚C/45min). These reaction conditions differ in sugar release between the two variants, reported previously by Bhagia et al.[20] Detailed description for the pretreatment conditions can be found in the Materials and Method section. SANS profiles were recorded at intervals during heating, retention and cooling phases (Figure 5.6 Panels A – C) to measure the structural changes in the biomass as they occurred during pretreatment (Figure 5.5). The SANS profiles were analyzed using the Unified Fit approach[31] using three structural levels that covered the Q range from $0.1\ \text{Å}^{-1} < Q < 0.4\ \text{Å}^{-1}$ (primary structural level), $0.01\ \text{Å}^{-1} < Q < 0.1\ \text{Å}^{-1}$ (secondary structural level) and $0.003 < Q < 0.01\ \text{Å}^{-1}$ (tertiary structural level). The fitting parameters obtained for each SANS profile at the different stages of the pretreatment reactions are shown in Figure 5.5 (Panels D-L) and Table S5.3.

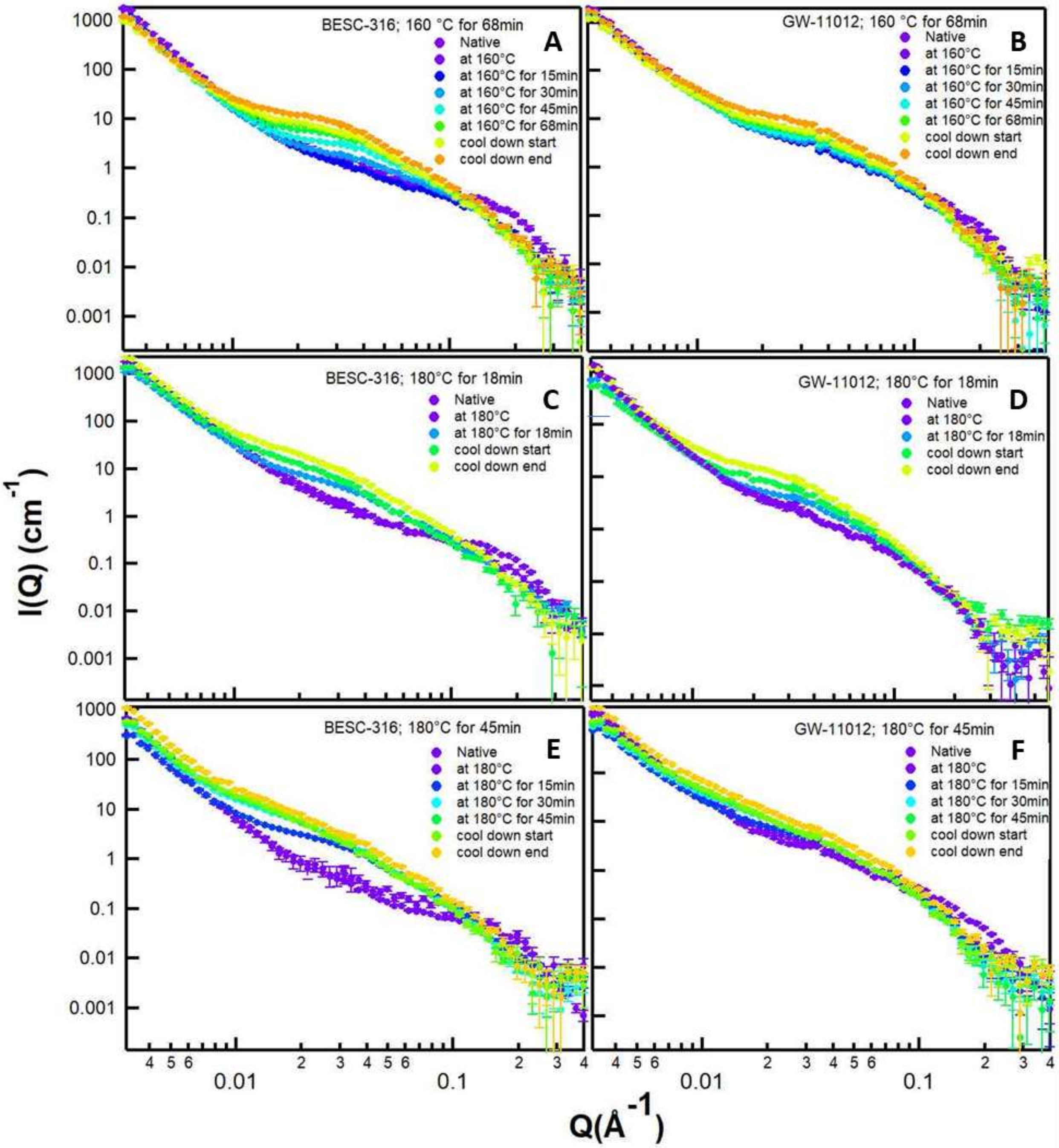


Figure 5.5: Time-resolved SANS data of hot water pretreatment of poplar variants BESC-316 and GW-11012. Pretreatments were performed in a pressure cell at 160˚C and 180˚C for the different retention times indicated in the panels A–F. The rainbow color sequence (violet to yellow) indicates the time intervals for the recorded SANS profiles.

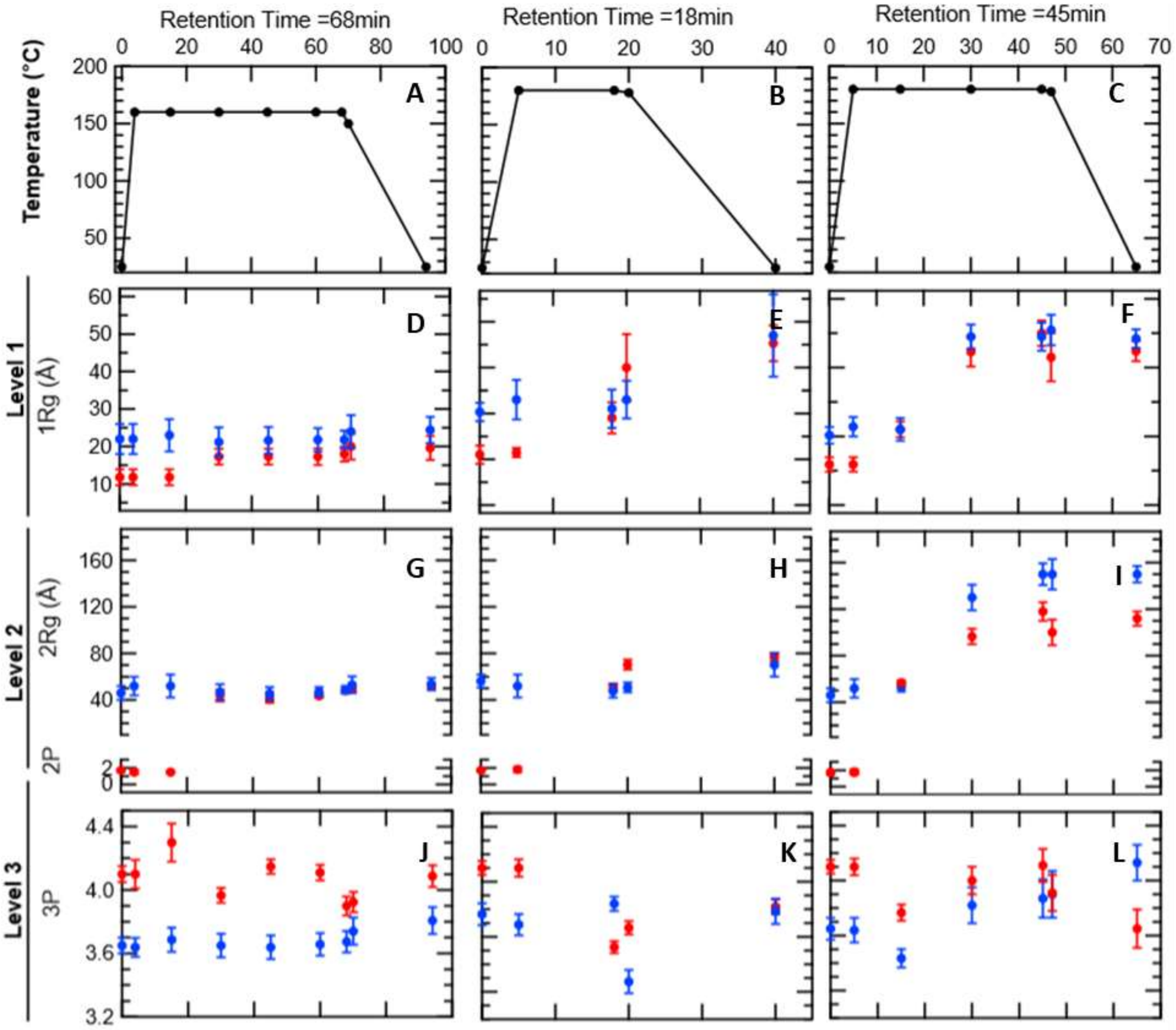


Figure 5.6: Structural parameters obtained from the analysis of *in-situ* SANS studies of hot water pretreatment reactions. *Panels A-C:* Temperature and time profiles for the pretreatment reactions. Closed circles represent the time points for the SANS profiles presented in Figure 4. *Panels D-L:* Plots of power-law exponents and R_g values obtained from the analysis of the SANS profiles at different time points for BESC-316 (red dots) and GW-11012 (blue dots) during the pretreatment reactions. Levels 1 – 3 represent the Q-ranges used in analysis of the SANS profiles, as described in the text. Selected fits and fit values are presented in the supplementary information.

Analysis of the structural changes in BESC-316 during the 160˚C/68min pretreatment shows the level-1 R_g, attributed to the cellulose microfibril cross-section dimension,[31] increases from 11 ± 2 Å to 19.6 ± 3.2 Å, indicating that the cellulose transitions from individual microfibrils to aggregated or coalesced microfibrils. For GW-11012, the level-1 R_g values are relatively unchanged before and after pretreatment, with values of 20.3 ± 2.4 Å and 24.4 ± 3.5 Å, respectively. This shows that the hot water pretreatment reaction had no impact on the already coalesced cellulose microfibrils of native GW-11012.

In the initial stage of the pretreatment, the 2nd structural level for BESC-316, is fit with a power law that transitions to a size feature as the pretreatment progresses. The appearance of this feature in the scattering curve was first observed about 30 min into LHW pretreatment at 160 ˚C. This feature was fit to a particle size of R_g = 44 Å that eventually increased to 52 Å for longer reaction times. We attribute this feature to the formation of lignin aggregates because delignification of the pretreated sample resulted in the disappearance of the particle size feature from this region of the SANS profile (Figure S4). This interpretation is consistent with previously published pretreatment studies with SANS.[30, 33, 39] Interestingly, as the pretreatment progressed through the retention phase and entered the cooling phase, the size of the lignin aggregate (~ 52 Å) did not increase but the intensity of the scattering feature continued to increase. This was captured numerically by the scalar value (G) in the Guinier part of the Unified fit equation which increased from 2.1 to 15 over the course of the pretreatment reaction (Table S3). The increase in G is interpreted as an increase in the number of lignin aggregates that were formed during the pretreatment reaction. The distribution of lignin in the cell wall of native GW-11012 is different from BESC-316, as described above. The initial size of the lignin aggregates was ~56 Å and didn’t change significantly during the course of the pretreatment reaction (Table S3). Similar to BESC- 316, the G value increased from 4 to 13 indicating an increase in the number of lignin aggregates formed during the pretreatment reaction.

Time resolved changes in the structure of BESC-316 and GW-11012 were also obtained for the LHW pretreatment conducted at 180˚C/18min (Figure 5.5 panel C and D). The cellulose microfibril cross section increased to 19 ± 3 Å (Level-1 R_g) after the 18 min retention time for BESC-316 indicating coalescence of the microfibrils and remained relatively unchanged for already aggregated GW-11012 cellulose microfibrils. Interestingly, the R_g value continued to increase past the retention phase to a final value R_g = 35 ± 4 Å in BESC-316 and R_g = 37 ± 9 Å in GW-11012 at the end of the pretreatment reaction. In order to determine if this increase was related to cellulose coalescence or another factor, we measured delignified BESC-316 after the pretreatment reaction was completed and obtained R_g of 23.0 ± 0.3 Å. This suggests that the additional intensity observed at Q ~ 0.15 $Å^{-1}$ during the pretreatment reaction was in fact due to the formation of lignin aggregates. The secondary regime followed a similar progression of structural changes to those observed for 160˚C/68min. The increased intensity in this region was fit with a Guinier function for both variants to yield R_g values of 77 ± 4 Å and 70 ± 10 Å for BESC-316 and GW-11012, respectively, at the end of the pretreatment reaction (Table S4). As before, this feature was not visible in the delignified curves after pretreatment and hence was assigned to lignin aggregates.

A similar trend in structural changes in the cell walls of the poplar variants were observed during the 180˚C/45min pretreatment reaction. The major differences were that lignin aggregate size (level-2 R_g) was larger, 150 ± 7 Å and 112 ± 6 Å for BESC-316 and GW-11012, respectively, compared to approximately 70 Å in the 180˚C/18min pretreatment reaction. In addition to the larger aggregate size, at 180 ˚C/45 min, there are also a greater number of lignin aggregates formed by the end of the pretreatment reaction (G~42 for BESC-316 and G~65 for GW-11012).

The 3rd structural level was fit using a power-law function. Changes in this parameter are interpreted as changes in the surface morphology of the cell wall. For instance, after pretreatment at 160˚C/68min, GW-11012 (3P ~ 3.81 ± 0.08) has a smoother surface morphology than in its native state but it remains rough compared to the pretreated BESC-316 (3P ~ 4.09 ± 0.07). Surface roughness has previously been related to greater cellulose accessibility.[31] The power-law values can be interpreted as increased cellulose accessible surface area after the 160˚C/68min pretreatment for GW-11012 compared to BESC-316. After the 180˚C/18min pretreatment, the exponent α for BESC-316 is 3.81 ± 0.05 similar to that obtained for GW-11012 (3P or $\alpha \sim 3.8 \pm 0.1$). For the 180˚C/45min, at which most difference in glucan release was observed,[20] the cell wall surface morphology for BESC-316 and GW-11012 were 4.17 ± 0.12 and 4.13 ± 0.14, respectively, indicating a smooth surface. It is important to note that the accuracy in the level-3 exponent (α) was low for both the samples because of the limited Q range available for the level-3 power-law fit.

5.4.2). Atomic structural differences between native and pretreated BESC-316 and GW-11012

X-ray diffraction (XRD) analysis was performed for the native samples and also the pretreated samples that were recovered after the *in-situ* SANS study described above. The 1D profiles for the native and pretreated samples shown in Figure 6 were derived from the equatorial scattering of the 2D pattern (Figure S5). Both BESC-316 and GW-11012 displayed the typical diffraction pattern for plant cellulose (I_β) consisting of a convoluted peak formed from two reflections of at (1-10) and (110) and a distinct peak at (200).[30]

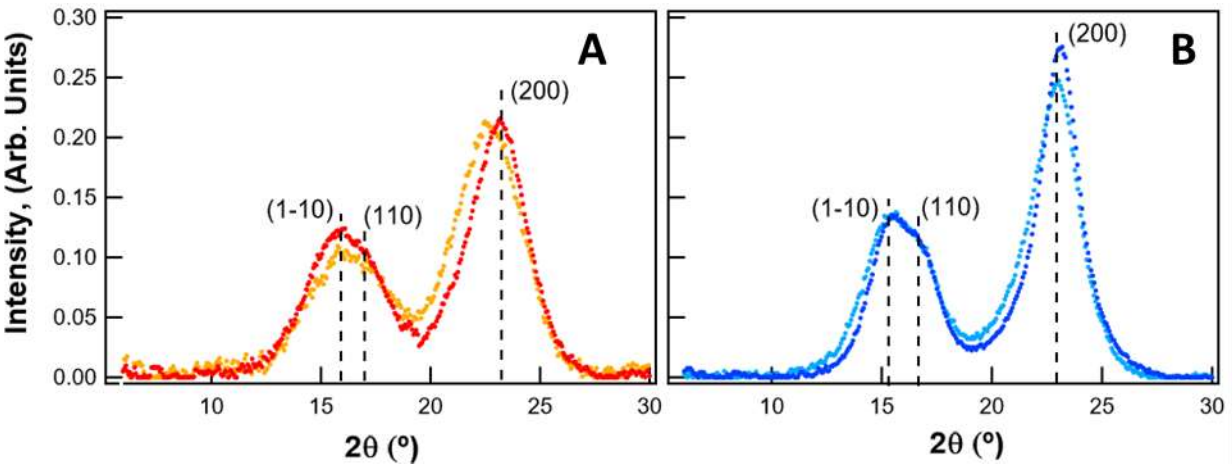


Figure 5.7: X-ray diffraction profiles of native BESC-316 (panel A) and GW-11012 (panel B) and after pretreatment at 180 ˚C for 45min. The curves were used to determine percent crystallinity and crystallite size in the (200) plane.

Table 5.3: Crystallinity, crystallite size and orientation parameter estimation. The crystallite size and orientation parameter were determined from the reflection at (200)

	Condition	Crystallite size (Å)	Crystallinity (%)	Orientation parameter (*f*)
BESC-316	native	26.0 ± 0.4	29	-0.1
	160˚C/68min	29.3 ± 0.3	33	-0.13
	180˚C/18min	29.7 ± 0.5	32	-0.14
	180˚C/45min	29.2 ± 0.4	34	-0.16
GW-11012	native	36.3 ± 0.6	41	-0.17
	160˚C/68min	36.2 ± 0.6	40.4	-0.23
	180˚C/18min	38.2 ± 0.3	45	-0.20
	180˚C/45min	37.8 ± 0.2	48	-0.30

5.3.2.1). Crystallite size and Crystallinity

The crystalline content and crystallite size (L_{200}) from native and pretreated BESC-316 and GW-11012 were calculated as described in the Materials and Methods (Table 3). Both the crystallinity and crystallite size are higher, ~1.4 fold, for native GW-11012 compared to BESC-316. As the pretreatment severity is increased, these values also increase indicating that the cellulose microfibrils become more crystalline and coalesce as a consequence of the pretreatment reaction.[41]

5.3.2.2). Crystallite Orientation

The 2D XRD pattern of the native GW-11012 exhibit significantly sharper reflections than native BESC-316 indicating a higher degree of cellulose crystallite orientation (Figure S5). After pretreatment, there is an increase in the number of visible reflections as well as a decrease in the arc size for both variants. This indicates that the crystallite orientation with respect to the fiber axis (c-axis) is further increased for both variants as a consequence of pretreatment.

The Hermann's orientation function [43-44] was used to quantify the degree of orientation of the cellulose crystallites with respect to the fiber axis to obtain the orientation factor (f). The (200) reflection was chosen to calculate f, because this reflection is strongest in both GW-11012 and BESC 316 and doesn't overlap with other reflections. The f_{200} value of -0.5 indicates that the crystallite orientation along the (200) plane is perpendicular to the fiber axis whereas a f_{200} of 0 indicates a random orientation of crystallites.[43-44] In this study, the crystallite dimension along the (200) plane shows very little orientation for native BESC-316 (-0.1) while the crystallites for native GW-11012 (-0.17) are comparably better oriented with respect to the fiber axis. After the pretreatment reactions, the proportional increase in the f values for both BESC-316 and GW-11012 are similar, indicating an increase in alignment of the crystallites along the fiber axis in both systems. However, after the most severe pretreatment condition, 180 ˚C for 45min, the cellulose crystallite orientation in GW11012 was approximately 1.9-fold greater compared to BESC-316.

5.4). Discussion

Our overall aim was to determine a structural basis for differences in sugar release between two naturally occurring poplars, BESC-316 and GW-11012, that were reported to have significant differences in lignin content in their cell walls.[20] Previous studies focused on characterizing the content and composition of the cell wall polymers of the natural poplar variants and attributed the changes observed in glucan release to lignin content and composition.[18, 20] Here, we used SANS to characterize cell wall structure at the nanoscale (1 – 100 nm) in native and pretreated lignocellulose, complemented by XRD to study the atomic structure of the crystallites within the cellulose microfibrils. Although the cellulose microfibrils were well-aligned with respect to the growth direction in both variants, SANS showed distinct differences in their packing and arrangement. In BESC-316, a defined microfibrillar distance (~28 Å) between neighboring microfibrils was evident, as is typically observed in SANS profiles of poplar secondary cell walls, [31, 45] indicating that the individual microfibrils are spatially separated and correlated with each other. In contrast, the SANS profiles of GW-11012 showed that the cellulose microfibrils had coalesced to form larger cellulose aggregates, a feature usually not associated with native poplar but has been observed for poplar grown under tension stress.[24] X-ray diffraction analysis also showed that the crystallite size along the (200) plane is larger for GW-11012 compared to BESC-316 supporting that neighboring GW-11012 microfibrils are aggregated. The mid-Q range SANS data provides information about matrix copolymer distribution, and for BESC-316 the data is similar to previously reported work. It can be interpreted as an inter-dispersed network of lignin and hemicellulose. [30-31, 33-34, 39] However, there was a clear difference in this Q-region for GW11012 that could be assigned to the presence of lignin aggregates in the cell wall. Although, lignin aggregates have not been not been observed previously in native poplar cell walls, uneven lignin distribution have been observed in the tension side of tension wood using coherence anti-stroke Raman scattering. [24] Lignin has previously been shown to coat the cellulose fibers and differences in its distribution can affect cellulose accessibility to enzymes. [13] Interestingly, measurement of the cellulose accessible surface area of the two variants showed that it is greater for GW11012 compared to BESC-316 and the amount of sugar released from enzymatic digestion of the native plants was also greater for GW11012.[20] [28] Overall, SANS analysis shows that reduced lignin content can a significant rearrangement of both cellulose microfibril organization and also lignin distribution in secondary cell walls of poplar. Lignin aggregation in thermochemically pretreated lignocellulose has previously shown to increase cellulose accessible surface area and its digestibly and explains the observation that sugar release from enzymatic digestion of the native plants was greater for GW11012. [20] [28]

As described in the introduction, LHW pretreatment of GW11012 performed at severity factors (log R_0) of 3.6 (160˚C/68min and 180˚C/18min) and $R_0 = 4$ (180˚C/45min) resulted in a corresponding increase in the measured sugar release. In contrast, the amount of sugar released from BESC-316 did not show the same trend as GW-11012 and was similar irrespective of the pretreatment severity.[20] By performing time-resolved *in situ* SANS measurements it was possible to observe the structural changes in the biomass as they occurred during the pretreatment reactions rather than just the initial and final states of the pretreated biomass. Comparison of the structural changes in GW-11012 and BESC-316 during the 160˚C/68min pretreatment reaction showed that the cellulose microfibrils in BESC-316, which are spatially separated in the native plant become aggregated, whereas there is little change observed in the already aggregated GW11012 microfibrils. The formation of the lignin aggregates in the two variants are similar; they start to a appear and grow in size to 41 Å ± 3 within 30 min, after which their size only increases slightly to 48 Å ± 2 by the end of the 68 min retention phase. However, the number of aggregates continues to increase as the pretreatment reaction progresses, to approximately 4-fold, from the start to the finish of the retention phase. The observation that the lignin aggregates do not grow in size but only in number during the retention phase suggests that they are spatially confined within the cell wall and a possible mechanism for their formation is by condensation of neighboring lignin molecules. [13, 46] Although the severity factor of the 180˚C/18 min and 160˚C/68min pretreatment condition was the same, the rate of formation of lignin aggregates was 2-fold faster and R_g of the aggregates were ~20% larger. This pretreatment temperature is well above the glass transition temperature of lignin (T_g = 80 – 140°C).[47] It is likely that the lignin becomes mobile in the cell wall allowing it to redistribute and form larger aggregates and suggests that higher a pretreatment temperature for a shorter time is more efficient. This trend in growth of lignin aggregate size and number is continued in the most severe pretreatment condition at 180˚C/45min, that produced lignin aggregates approximately 25% larger and a greater in number in GW-11012 compared to BESC-316. We can conclude that based on time-resolved *in situ* SANS analysis, the temporal changes observed during the different pretreatment regimes in the cell walls of the two variants all followed a similar trend. Additionally, comparison of the final pretreated cell wall structure after the most severe pretreatment, showed no significant differences that would explain the increased sugar release in GW-11012 that was observed. This suggests that the nanoscale organization of the pretreated cell walls, primarily lignin aggregation, is not the main contributing factor to its overall susceptibility to enzymatic digestibility.

Cellulose accessibility was also analyzed using the Simon's Stain [48] to determine if there were differences between BESC-316 and GW-11012 after the different pretreatment regimes that would give rise to the observed differences in sugar release (Table S1). There was significant increase in the cellulose accessibility after the 160°C/68min pretreatment reaction for both variants. However, no clear trend was evident in the increase in these values for the pretreatments at 180°C except that after the 180°C/45min pretreatment the cellulose accessibility values were similar for both variants. These results are consistent with the fitting parameters obtained from the 3rd level of the Unified Fit analysis. As described in the Results section, the exponent value for level-3 obtained from the fit is interpreted as measure of cell wall roughness.[31] The values obtained for GW-11012 and BESC-316 after the180˚C/45min pretreatment were 4.13 +/- 0.14 and 4.17 +/- 0.12, respectively. Together, these data suggest that there are reasons other than cellulose accessibility and cell wall roughness that are responsible for greater sugar release from pretreated GW-11012. The presence of acetic acid as an inhibitory product,[7] that can be formed due to deacetylation of hemicelluloses during LHW pretreatment, was also discounted as a reason for the differences in sugar release. There was no difference in the amount present in the pretreatment liquors of the two variants (Table S2). Also, the potential effects of other inhibitory products such as 5-Hydroxymethylfurfural (5-HMF) and furfural was addressed by Bhagia et al. [20]

Cellulose structure is known to influence enzymatic hydrolysis.[6] The atomic scale differences in the cellulose structure of BESC-316 and GW-11012 including crystallite size, crystallinity and crystallite orientation were quantified. These values were all significantly higher in GW-11012 compared to BESC-316 indicating a higher degree of organization of cellulose crystallites in this variant, and these differences persisted after the pretreatment. Comparison of changes that occurred to the individual variants showed only a slight increase in cellulose crystallite size and crystallinity after pretreatment (regardless of severity) as compared to the native samples. Similar changes in cellulose structure after dilute acid and LHW pretreatments have been reported previously.[49] The most notable difference in cellulose structure between the variants was the orientation of the crystallites in the cellulose microfibrils. As described in the results section, the crystallites in native GW-11012 cellulose were significantly better oriented compared to BESC-316. A likely reason for this is that the difference in lignin distribution in the two variants allows the cellulose microfibrils to coalesce in GW-11012 and this in turn increases H-bonding interactions between adjacent microfibrils resulting in better aligned microfibrils with increased orientation of crystallites. Pretreatment further increased the alignment of crystallites in both BESC-316 and GW-11012 and the orientation was approximately 2-fold greater in both the variants after pretreatment when compared to the native samples. However, the absolute orientation of the crystallites in GW11012 was approximately twice that of BESC-316 after the most severe pretreatment.

A consequence of having better aligned crystallites relates to the digestibility of crystalline cellulose. Processive cellobiohydrolases are a class of cellulases that bind crystalline cellulose via a cellulose binding module and move along the cellulose surface as they detach and hydrolyze cellulose chains releasing cellobiose. Cellulase processivity been observed on cellulose surfaces using fluorescently labelled cellulose and demonstrated higher cellobiose greater release from crystalline cellulose compared to amorphous cellulose.[50] A similar approach using atomic force microcopy was able to track the rate that cellulases moved along cellulose surfaces and how long they remained attached to the microfibrils.[51-52] This study also showed increased attachment to crystalline cellulose compared to amorphous cellulose. The movement of cellulases along the cellulose surface has also been shown to be preferential along a certain cellulose crystallite phase.[51-52] The implication for this, in the context of the present study, is that the amount of sugar released from GW11012 after the most severe pretreatment condition was 1.6 times compared to BESC-316 but there were no significant differences in the nanoscale (1 – 100 nm) structure observed. The most significant difference between the two variants that would explain the increase in sugar release is the orientation of the crystallites in the microfibrils. Based on this, we can propose that the increased crystallite alignment in GW11012 improves the efficiency of processive cellobiohydrolases by providing an obstacle free path for the enzymes allowing them to bind and remain attached to the cellulose chains longer resulting in higher glucan release.

5.5). Conclusion

Our ability to investigate cellulose, hemicellulose and lignin interactions in the plant cell wall is challenging due to the highly interconnected and heterogeneous lignocellulose structure. The establishment of a large collection of geographically distributed undomesticated *P. trichocarpa* genotypes provided a way to perform parametric studies that relate cell wall composition to recalcitrance and demonstrated that a lower lignin content and composition with an S/G ratio greater than two, were indicators of reduced recalcitrance to enzymatic digestion.[18] Here, we reported a multiscale structural analysis of a high and low lignin genotype from this population to determine if there was a structural basis for the difference in sugar release from these genotypes. Our analysis showed clear differences in the arrangement of the cell wall polymers in the low lignin variant at the nanoscale using SANS. Specifically, the cellulose microfibrils were coalesced rather than spatially separated as is typically observed in the secondary cell walls of poplar. In addition, lignin was aggregated rather than homogenously interspersed with other matrix co-polymers. However, after subjecting each genotype to LHW pretreatment regimes of increasing severity, it was observed that the chronology of structural changes that took place in the variants was similar and the final structure of the pretreated lignocelluloses after the most severe pretreatment were almost indistinguishable. It was clear from these data that the nanoscale structural features of the pretreated samples could not adequately explain the 1.6 times increase in sugar release observed in GW-11012 compared to BESC 316, the high lignin variant. Atomic scale structural analysis of the cellulose microfibrils revealed differences between the two variants. The higher crystallinity and larger crystallite size in native and pretreated GW-11012 are likely because of the close association between adjacent microfibrils. However, the most significant difference was the orientation of the crystallites in the microfibrils. This was approximately 2-fold higher in the low lignin variant compared to the high lignin variant. Given that cellulose crystallite orientation was a consistent difference between the two variants before and after pretreatment, we can speculate that the higher degree of organization within GW-11012 microfibrils provides a cellulase accessible surface favorable for binding and movement of processive glycoside hydrolases along cellulose microfibril surfaces. Overall, this study provides new and unexpected insights into the interplay of plant cell wall polymers and their effect on biomass recalcitrance.

Acknowledgments

R. S., S.B, S. V. P., A.J.R, H. O'N., and B.H.D. acknowledge the support of the Genomic Science Program, Office of Biological and Environmental Research (OBER), U. S. Department of Energy (DOE), under Contract FWP ERKP752 for this research. Thermochemical pretreatment and XRD measurements were performed at the Center for Nanophase Materials Sciences (CNMS) supported by Office of Basic Energy Sciences (BES), DOE. The SANS studies on Bio-SANS were supported by the OBER funded Center for Structural Molecular Biology (CSMB) under Contract FWP ERKP291, using the High Flux Isotope Reactor supported by the BES, DOE. This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S.

Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan)

ASSOCIATED CONTENT Supplementary file 1

AUTHOR INFORMATION

*Corresponding authors: oneillhm@ornl.gov‡, Present address: Neutron Scattering Division, Oak Ridge National Laboratory, TN:37830: davisonbh@ornl.gov Present address: BioSciences Division, Oak Ridge National Laboratory, TN:37830.

All authors have given approval to the final version of the manuscript.

5.7). Supplementary information

5.7.1). Preparation and pretreatment of ground biomass samples

The debarked stem segments were air dried to a constant weight and knife milled through a 20-mesh screen. Previously, the milled poplar wood was processed with high throughput hot water pretreatment and co- hydrolyzed (HTPH) and the amount of glucan release was quantified. The hot water pretreatment processing conditions included no pretreatment, 160°C/68 min, 180°C/18min and 180°C/45min.[20] Table 5.4 shows the previously reported glucan release from these two poplars after the different processing conditions.

For liquid hot water pretreatments, the ground samples were processed by first soaking 2.12g of material overnight in 40 ml of distilled water. The presoaked slurry was transferred to a 75 mL total capacity Parr reactor with a 5% dry solid loading. The reactor was sealed and heated to 160

˚C or 180 ˚C. The system took ~18min to reach 160 °C and ~25 min to reach 180 ˚C. For the pretreatments carried at 160 ˚C the retention time was 68 min while for the pretreatment at 180 ˚C, the retention time was either for 18 min or 45 min. The reaction was quenched by placing the sealed Parr reactor on ice until the temperature reached 25 ˚C. The pretreated wood slurry was recovered and washed thoroughly with distilled water. The pretreatment liquor was also recovered and neutralized to pH ~6.0. Acetic acid content in the neutralized pretreated liquor was determined using high performance anion exchange chromatography with UV-Vis detection.

5.7.2). Quantitative analysis of pretreatment liquor for sugars, inhibitors and acetic acid

The cellulose, hemicellulose and lignin content were determined using the previously reported standard methods by averaging three replicates of milled samples for each variant.[28] Acetic acid is expected to be found in the pretreatment liquor if the hemicellulose sugar groups are acetylated. The pretreatment liquor from the ground slurry was collected and the insoluble fraction was washed with water to ensure the pretreatment liquor was completely removed. The liquor and wash were analyzed for acetic acid content. Both BESC-316 and GW-11012 released similar amounts of acetic acid into the pretreatment liquor regardless of the severity of the pretreatment (Table 5.5).

5.7.3). Cellulose accessibility measurements using Simons' stain

Sugar release by enzyme hydrolysis is dependent cellulose accessibility to enzymes. Simons' stain was used to determine the cellulose accessible surface area for the native and hot water pretreated wood, as previously described.[61]

Table 5.4: Glucan release from BESC-316 and GW-11012 after the different pretreatment conditions and co- hydrolysis[20]

Pretreatment condition	BESC-316 (mg/100 mg biomass)	GW-11012 (mg/100 mg biomass)
No pretreatment, enzymatic hydrolysis	12	25
HTPH, 160˚ C, 68.1 min, severity factor =3.6	52	53
HTPH, 180˚ C, 17.6 min, severity factor =3.6	60	72
HTPH, 180 ˚C, 44.1 min, severity factor =4.0	57	89

Table 5.5: Quantification of the acetic acid released into the pretreatment liquor from the pretreated biomass

Sample	Pretreatment	Acetic Acid (mg/g)
BESC-316	160˚C/ 68min	5
	180˚ C/18min	5
	180˚C/45min	15
GW-11012	160˚C/68min	4
	180˚ C/18min	8
	180˚C/45min	15

The maximum dye absorption to the biomass is shown in Table 5.6. The pretreated wood samples after the 160˚C/68min pretreatment bind significantly more dye than the native wood samples and as the pretreatment reactions become more severe the cellulose accessibility increases, as evidenced by increased dye binding to the cellulose. This trend is continued for the 180˚C/45min pretreatment, for which the maximum difference in sugar release has been reported.[18] Both poplar variants, bind a similar amount of dye indicating that the accessible surface area for BESC-316 and GW-11012 is the same after this pretreatment condition.

5.7.4). Small-angle neutron scattering data analysis

The unified fitting approach[32, 38] was used to fit all SANS curves (see materials and methods section in the main text). SANS curves for native BESC-316 and GW-11012 and the SANS curves after the different pretreatment regimes (160˚C/68min, 180˚C/18min and 180˚C/45min) are shown in Figure 5.8. The Q-range from 0.003 $Å^{-1}$ to 0.3 $Å^{-1}$ was used to fit the curves. In order to obtain the overall fit, three fitting levels were used and the fit parameters are shown in Table 5.7. Each fitting level was comprised of two fitting functions, a Guinier that gives information about particle size (R_g) and power law that gives information about the particle morphology. The fitting parameters were subjected to the uncertainty test to provide an estimate of error in each fitting parameter.

5.7.5). SANS analysis of delignified poplar variants

The 2D images and the 1D scattering curves of delignified BESC-316 and GW-11012 are shown in Figures 5.9 and 5.10, respectively. The Q region spanning from 0.04 to 0.2 $Å^{-1}$ in 1D profile of delignified BESC-316 and GW-11012 shows increase in intensity as compared to the native BESC-316 and GW-11012.

Table 5.6: Maximum dye adsorption for the native and pretreated BESC-316 and GW-11012

Pretreatment	Maximum dye adsorption (mg/g biomass)	
	BESC-316	GW-11012
Native[28]	26	40
160 ˚C/ 68 min	70	87
180 ˚C /18 min	92	104
180 ˚C/ 45 min	104	107

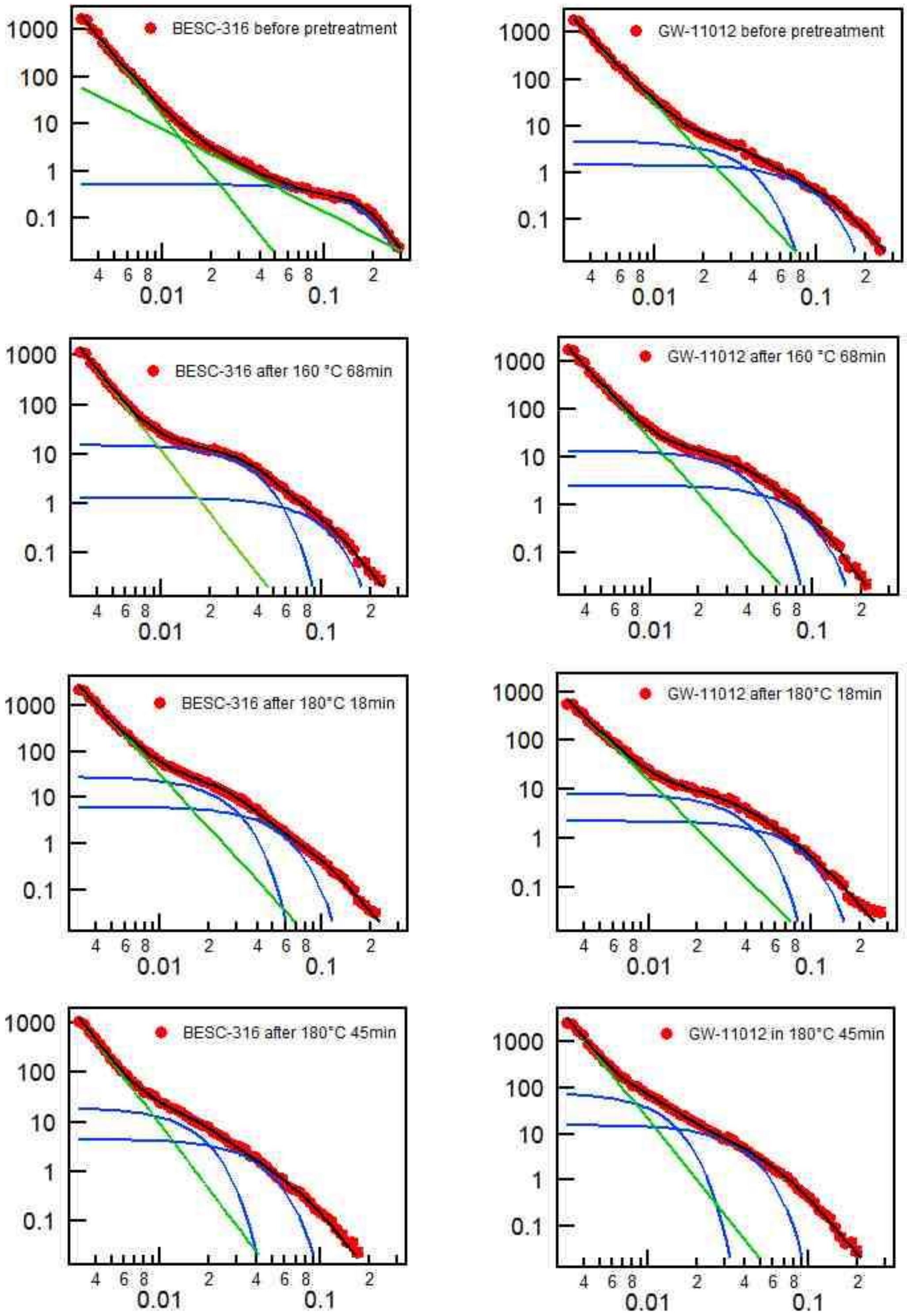


Figure 5.8: Unified Fits for selected scattering curves. Each plot has experimental scattering curve (red), unified fit (black) obtained by summing three levels. Shown in the plot is the Guinier fit (blue) for level1, Porod fit (green) for level 2 for BESC-316, Guinier fit (blue) for level 2 for rest of all samples and Porod fit (green) for level3.

Table 5.7: Fit parameters obtained from the Unified Fit Approach for the different pretreatment regimes

	Stages in pretreatment	**Level 1**		**Level 2**			**Level 3**
		1G	**1Rg**	**2G**	**2P**	**2Rg**	**3P**
BESC-316 160 ˚C/68min	25 ˚C		11 +/- 2		1.7+/- 0.19	**400**	4.1
	160 ˚C	0.4766	11.44 +/- 0.63	**61.36**	1.8 +/- 0.3	**400**	4.417 +/- 0.064
	160 ˚C for 15min	0.215	11.8 +/- 2.1	**11.62**	1.53 +/- 0.27	**235.6**	4.30 +/- 0.12
	160 ˚C for 30min	0.7005	17.3 +/- 2.1	2.121		44.0 +/- 5.2	3.966 +/- 0.047
	160 ˚C for 45min	0.7272	17.3 +/- 2.1	3.78		40.5 +/- 2.9	4.148 +/- 0.046
	160 ˚C for 60min	0.8046	17.3 +/- 2.2	6.435		43.6 +/- 1.9	4.110 +/- 0.049
	160 ˚C for 68min	0.9437	18 +/- 2	8.445		48.0 +/- 1.7	3.90 +/- 0.06
	cool down to 155 ˚C	1.155	*20.0 +/- 3.5*	9.438		49.9 +/- 2.5	3.924 +/- 0.063
	to 25 ˚C	1.281	19.6 +/- 3.2	14.95		*52.0 +/- 2.7*	4.087 +/- 0.069
		1G	**1Rg**	**2G**	**2P**	**2Rg**	**3P**
GW-11012 160 ˚C/68min	25 ˚C	1.34	20.3 +/- 2.4	4.901		56.2 +/- 5.5	3.764 +/- 0.079
	160 ˚C	1.423	23.0 +/- 4.3	3.59		52.0 +/- 9.9	3.687 +/- 0.077
	160 ˚C for 15min	1.423	21.2 +/- 3.9	3.59		46.6 +/- 6.8	3.650 +/- 0.075
	160 ˚C for 30min	1.379	21.6 +/- 3.6	4.83		45.5 +/- 5.5	3.639 +/- 0.075
	160 ˚C for 45min	1.697	21.8 +/- 3.1	6.102		47 +/- 4	3.657 +/- 0.071
	160 ˚C for 60min	1.679	21.8 +/- 2.4	7.658		48.7 +/- 3.3	3.674 +/- 0.068
	160 ˚C for 68min	2.325	24.0 +/- 4.4	8.279		53.1 +/- 7.1	3.740 +/- 0.086
	cool down to 155 ˚C	2.325	24.0 +/- 4.4	8.279		53.1 +/- 7.1	3.740 +/- 0.086
	to 25 ˚C	2.481	24.4 +/- 3.5	13.17		53.7 +/- 5.2	3.808 +/- 0.084

Table 5.7: (continued). Fit parameters obtained from the Unified Fit Approach for the different pretreatment regimes

	Stages in pretreatment	**Level 1**		**Level 2**			**Level 3**
		1G	**1Rg**	**2G**		**1G**	**1Rg**
BESC-316 180 ˚C/18min	25 ˚C		11 +/- 2		1.7+/- 0.19	**400**	4.1
	180 ˚C	0.4766	11.44 +/- 0.63	**61.36**	1.8 +/- 0.3	**400**	4.417 +/- 0.064
	180 ˚C for 15min	0.9887	19.0 +/- 3.4	6.735		51 +/- 3	3.524 +/- 0.044
	cool down to 175 ˚C	2.475	*30.0 +/- 7.3*	16.16		70.4 +/- 4.3	3.665 +/- 0.049
	to 25 ˚C	6.331	35.3 +/- 3.9	27.93		76.9 +/- 3.6	3.818 +/- 0.053
		1G	**1Rg**	**2G**	**2P**	**2Rg**	**3P**
GW-11012 180 ˚C/18min	25 ˚C	1.34	20.3 +/- 2.4	4.901		56.2 +/- 5.5	3.764 +/- 0.079
	180 ˚C	1.423	23.0 +/- 4.3	3.59		52.0 +/- 9.9	3.687 +/- 0.077
	180 ˚C for 15min	1.402	21.0 +/- 4.2	4.311		47.6 +/- 5.7	3.840 +/- 0.051
	cool down to 175 ˚C	2.239	23.0 +/- 4.1	8.135		50.7 +/- 4.2	3.273 +/- 0.084
	to 25 ˚C	9.044	37 +/- 9	11.93		70 +/- 10	3.784 +/- 0.092

Table 5.7: (continued). Fit parameters obtained from the Unified Fit Approach for the different pretreatment regimes

	Stages in pretreatment	**Level 1**		**Level 2**			**Level 3**
		1G	**1Rg**	**2G**		**1G**	**1Rg**
		1G	**1Rg**	**2G**		**2Rg**	**3P**
BESC-316 180 ˚C/45min	25		11 +/- 2		1.7+/- 0.19	**400**	4.1
	180 ˚C	0.5008	22.0 +/- 2.3	3.56		56.8 +/- 2.2	3.767 +/- 0.058
	180 ˚C for 15min	3.24	44.6 +/- 4.3	9.859		96.5 +/- 6.7	4.0 +/- 0.1
	180 ˚C for 30min	4.211	50.0 +/- 3.7	14.48		118 +/- 8	4.11 +/- 0.12
	180 ˚C for 45min	2.779	43 +/- 7	12.53		100 +/- 11	3.91 +/- 0.13
	cool down to 175 ˚C	4.374	44.8 +/- 4.3	19.04		110 +/- 11	4.27 +/- 0.12
	to 25 ˚C	8.325	44.2 +/- 2.9	41.85		112.4 +/- 6.1	4.17 +/- 0.12
		1G	**1Rg**	**2G**		**2Rg**	**3P**
GW-11012 180 ˚C/45min	25	1.34	20.3 +/- 2.4	4.901		56.2 +/- 5.5	3.764 +/- 0.079
	180 ˚C	1.575	22.8 +/- 2.8	7.269		52.8 +/- 3.6	3.434 +/- 0.067
	180 ˚C for 15min	9.778	48.0 +/- 3.3	23.12		130 +/- 11	3.82 +/- 0.13
	180 ˚C for 30min	9.87	49.0 +/- 3.5	35.92		140.5 +/- 9.3	3.87 +/- 0.14
	180 ˚C for 45min	10.22	49.0 +/- 4.1	42.77		150 +/- 13	3.90 +/- 0.17
	cool down to 175 ˚C	16.46	50.9 +/- 4.4	80.64		160 +/- 12	4.37 +/- 0.13
	to 25 ˚C	12.84	48.4 +/- 2.7	64.84		150 +/- 7	4.13 +/- 0.14

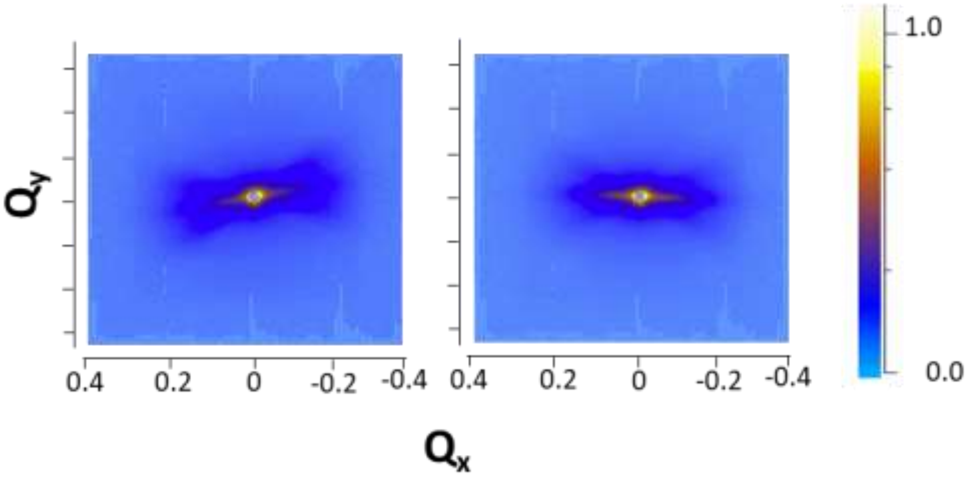


Figure 5.9: Two-dimensional detector intensity images are plotted as a function of Q_x vs. Q_y obtained measuring delignified BESC-316 (left) and GW-11012 (right) measured in 100% D_2O. The measurements were made at the EQ-SANS beamline of SNS at the ORNL

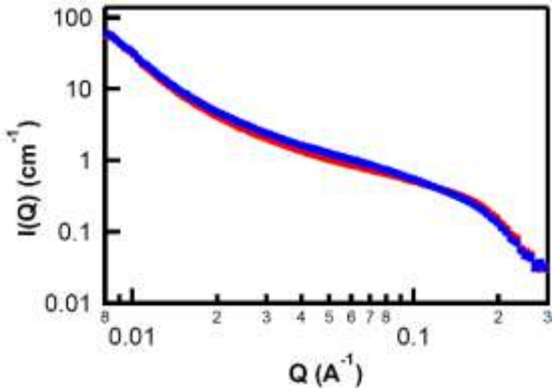


Figure 5.10: The SANS scattering 1D profile obtained by plotting of I(Q) vs Q for the delignified native BESC- 316 (red) and GW-11012 (blue)

A previous report investigated balsa wood using scanning electron microscopy showed that removal of lignin created pores in the cell walls which were quantified using Brunauer-Emmett-Teller (BET) surface area analysis. A board distribution in the pore size was observed using BET. Additionally, SAXS analysis of the same delignified wood samples showed an increase in scattering intensity at 0.13 Å^{-1} that was attributed to the presence of pores of size 4.8 nm ($using\ q = \frac{2\pi}{d}$).[61] Hence the increase in intensity in the Q range from 0.04 - 0.2 Å^{-1} could be due to the formation of pores after lignin removal. The SANS profiles were fit using the unified fitting approach (Table 5.8).

5.7.6). Comparison of pretreated wood before and after delignification

5.7.7). 2D X-ray diffraction images for the wood samples

The 2D images of the native and pretreated wood are shown in Figure 5.12. The images were corrected for the air scattering and imaging plate background. The reflection arcs are sharper in the images for native GW-11012 compared to BESC-316. For both GW-11012 and BESC-316, the reflections become sharper as a function of increase in pretreatment severity. However, the pretreated GW-11012 has much more sharper reflections compared to BESC-316 indicating a higher degree of orientation that was quantified using the Hermann's orientation factor (see main text). The diffraction patterns were converted to 1D patterns and from them the crystallinity, and crystallite size were determined. The description to calculate these parameters ig given in the materials and method section.

Table 5.8: Fit parameters obtained for delignified BESC-316 and GW-11012 using the Unified Fit

Delignified Sample	**Level1***	**Level2***
	R_g (Å)	**2P**
BESC-316	10.0 ± 0.3	1.40 ± 0.11
GW-11012	10.5 ± 0.3	1.25 ± 0.08

* The range of fitting for level-1 was 0.1 Å^{-1} <Q< 0.3 Å^{-1} and for level-2 was 0.002 Å^{-1} <Q<0.1 Å^{-1}

.

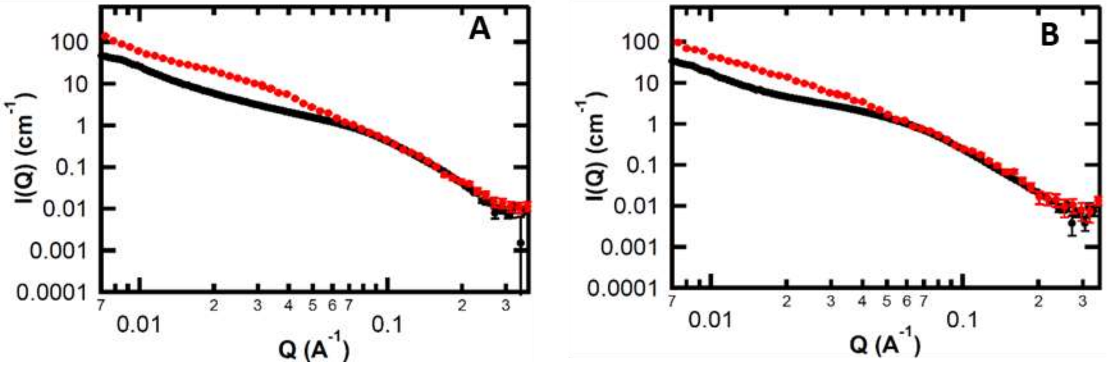


Figure 5.11: Comparison of pretreated BESC-316 before and after delignification. Panels A and B are for 180˚C/18min and 180˚C/45min, respectively. Red and black dots are for SANS data obtained before and after delignification of the pretreated sample, respectively.

Table 5.9: Fit parameters obtained for delignified pretreated BESC-316 at 180˚C/18min and 180˚C/45min

Delignified Sample	Pretreatment condition	Level 1*	Level 2*
		R_g (Å)	P
BESC-316	180 ˚C/18min	23.0 ± 0.3	2.4 ± 0.04
BESC-316	180 ˚C/45min	31.0 ± 0.4	2.5 ± 0.1

*The range of fitting for level 1 was 0.1 $Å^{-1}$ <Q< 0.3 $Å^{-1}$ and for level 2 was 0.002 $Å^{-1}$ <Q<0.1 $Å^{-1}$

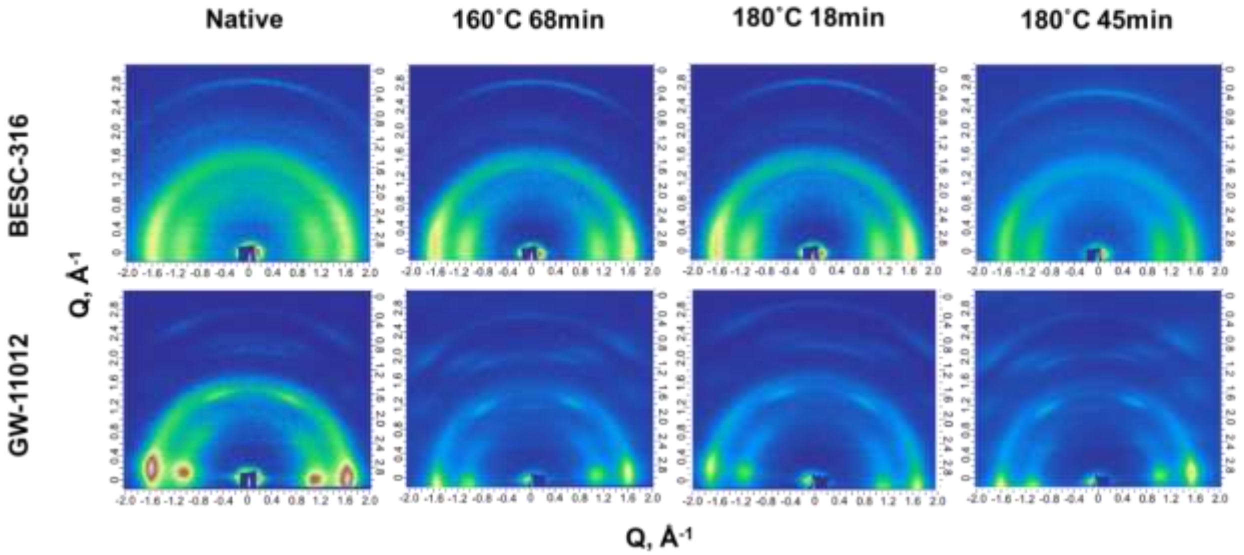


Figure 5.12: Background subtracted XRD pattern for BESC-316 and GW-11012

5.6). References

1. Balan, V.; Chiaramonti, D.; Kumar, S., Review of US and EU initiatives toward development, demonstration, and commercialization of lignocellulosic biofuels. *Biofuels, Bioproducts and Biorefining* **2013,** *7* (6), 732-759.
2. Ragauskas, A. J.; Williams, C. K.; Davison, B. H.; Britovsek, G.; Cairney, J.; Eckert, C. A.; Frederick, W. J.; Hallett, J. P.; Leak, D. J.; Liotta, C. L.; Mielenz, J. R.; Murphy, R.; Templer, R.; Tschaplinski, T., The Path Forward for Biofuels and Biomaterials. *Science* **2006,** *311* (5760), 484-489.
3. Lynd, L. R.; Liang, X.; Biddy, M. J.; Allee, A.; Cai, H.; Foust, T.; Himmel, M. E.; Laser, M. S.; Wang, M.; Wyman, C. E., Cellulosic ethanol: status and innovation. **2017,** (1879-0429 (Electronic)).
4. Meng, X.; Ragauskas, A. J., Recent advances in understanding the role of cellulose accessibility in enzymatic hydrolysis of lignocellulosic substrates. *Current Opinion in Biotechnology* **2014,** *27*, 150-158.
5. Li, M.; Pu, Y.; Ragauskas, A. J., Current Understanding of the Correlation of Lignin Structure with Biomass Recalcitrance. *Front Chem* **2016,** *4*, 45-45.
6. Yang, B.; Dai, Z.; Ding, S.-Y.; Wyman, C. E., Enzymatic hydrolysis of cellulosic biomass. *Biofuels* **2011,** *2* (4), 421-449.
7. Johnson, A. M.; Kim, H.; Ralph, J.; Mansfield, S. D., Natural acetylation impacts carbohydrate recovery during deconstruction of Populus trichocarpa wood. *Biotechnology for Biofuels* **2017,** *10* (1), 48.
8. Chum, H. L.; Johnson, D. K.; Black, S.; Baker, J.; Grohmann, K.; Sarkanen, K. V.; Wallace, K.; Schroeder, H. A., Organosolv pretreatment for enzymatic hydrolysis of poplars: I. Enzyme hydrolysis of cellulosic residues. *Biotechnol Bioeng* **1988,** *31* (7), 643-9.
9. DeMartini, J. D.; Pattathil, S.; Miller, J. S.; Li, H.; Hahn, M. G.; Wyman, C. E., Investigating plant cell wall components that affect biomass recalcitrance in poplar and switchgrass. *Energy & Environmental Science* **2013,** *6* (3), 898-909.
10. Ishizawa, C. I.; Jeoh, T.; Adney, W. S.; Himmel, M. E.; Johnson, D. K.; Davis, M. F., Can delignification decrease cellulose digestibility in acid pretreated corn stover? *Cellulose* **2009,** *16* (4), 677-686.
11. Lacayo, C. I.; Hwang Ms Fau - Ding, S.-Y.; Ding Sy Fau - Thelen, M. P.; Thelen, M. P., Lignin depletion enhances the digestibility of cellulose in cultured xylem cells. *PloS One* **2013,** *8* (7).
12. Himmel, M. E.; Ding, S.-Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D., Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* **2007,** *315* (5813), 804-807.
13. Donohoe, B. S.; Decker, S. R.; Tucker, M. P.; Himmel, M. E.; Vinzant, T. B., Visualizing lignin coalescence and migration through maize cell walls following thermochemical pretreatment. *Biotechnology and Bioengineering* **2008,** *101* (5), 913-925.
14. Rahikainen, J.; Mikander, S.; Marjamaa, K.; Tamminen, T.; Lappas, A.; Viikari, L.; Kruus, K., Inhibition of enzymatic hydrolysis by residual lignins from softwood—study of enzyme binding and inactivation on lignin-rich surface. *Biotechnology and Bioengineering* **2011,** *108* (12), 2823-2834.
15. Vanholme, R.; Morreel K Fau - Ralph, J.; Ralph J Fau - Boerjan, W.; Boerjan, W., Lignin engineering. *Current Opinion in Plant Biotechnology* **2008,** *11* (3), 278-85.

16. Davin, L. B.; Jourdes M Fau - Patten, A. M.; Patten Am Fau - Kim, K.-W.; Kim Kw Fau - Vassao, D. G.; Vassao Dg Fau - Lewis, N. G.; Lewis, N. G., Dissection of lignin macromolecular configuration and assembly: comparison to related biochemical processes in allyl/propenyl phenol and lignan biosynthesis. *Natural Products Report* **2008,** *25* (6), 1015-90.

17. Davison, B. H.; Drescher, S. R.; Tuskan, G. A.; Davis, M. F.; Nghiem, N. P., Variation of S/G ratio and lignin content in a Populus family influences the release of xylose by dilute acid hydrolysis. *Applied Biochemistry and Biotechnology* **2006,** *130* (1), 427-435.

18. Studer, M. H.; DeMartini, J. D.; Davis, M. F.; Sykes, R. W.; Davison, B.; Keller, M.; Tuskan, G. A.; Wyman, C. E., Lignin content in natural Populus variants affects sugar release. *Proceedings of the National Academy of Sciences* **2011,** *108* (15), 6300-6305.

19. Muchero, W.; Guo, J.; DiFazio, S. P.; Chen, J.-G.; Ranjan, P.; Slavov, G. T.; Gunter, L. E.; Jawdy, S.; Bryan, A. C.; Sykes, R.; Ziebell, A.; Klápště, J.; Porth, I.; Skyba, O.; Unda, F.; El-Kassaby, Y. A.; Douglas, C. J.; Mansfield, S. D.; Martin, J.; Schackwitz, W.; Evans, L. M.; Czarnecki, O.; Tuskan, G. A., High-resolution genetic mapping of allelic variants associated with cell wall chemistry in Populus. *BMC Genomics* **2015,** *16* (1), 24.

20. Bhagia, S.; Muchero, W.; Kumar, R.; Tuskan, G. A.; Wyman, C. E., Natural genetic variability reduces recalcitrance in poplar. *Biotechnology for Biofuels* **2016,** *9* (1), 106.

21. Zhao, Y.; Man, Y.; Wen, J.; Guo, Y.; Lin, J., Advances in Imaging Plant Cell Walls. *Trends in Plant Science* **2019**.

22. Beecher, J. F.; Hunt, C. G.; Zhu, J. Y., Tools for the Characterization of Biomass at the Nanometer Scale. **2009**, 61-90.

23. Zeng, Y.; Saar, B. G.; Friedrich, M. G.; Chen, F.; Liu, Y.-S.; Dixon, R. A.; Himmel, M. E.; Xie, X. S.; Ding, S.-Y., Imaging Lignin-Downregulated Alfalfa Using Coherent Anti-Stokes Raman Scattering Microscopy. *BioEnergy Research* **2010,** *3* (3), 272-277.

24. Foston, M.; Hubbell, C. A.; Samuel, R.; Jung, S.; Fan, H.; Ding, S.-Y.; Zeng, Y.; Jawdy, S.; Davis, M.; Sykes, R.; Gjersing, E.; Tuskan, G. A.; Kalluri, U.; Ragauskas, A. J., Chemical, ultrastructural and supramolecular analysis of tension wood in Populus tremula x alba as a model substrate for reduced recalcitrance. *Energy & Environmental Science* **2011,** *4* (12), 4962-4971.

25. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Pöppler, A.-C.; Brown, S. P.; deAzevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nature Communications* **2016,** *7* (1), 13902.

26. Terrett, O. M.; Dupree, P., Covalent interactions between lignin and hemicelluloses in plant secondary cell walls. *Current Opinion in Biotechnology* **2019,** *56*, 97-104.

27. Cheng, G.; Zhang, X.; Simmons, B.; Singh, S., Theory, practice and prospects of X-ray and neutron scattering for lignocellulosic biomass characterization: towards understanding biomass pretreatment. *Energy & Environmental Science* **2015,** *8* (2), 436-455.

28. Yoo, C. G.; Yang, Y.; Pu, Y.; Meng, X.; Muchero, W.; Yee, K. L.; Thompson, O. A.; Rodriguez, M.; Bali, G.; Engle, N. L.; Lindquist, E.; Singan, V.; Schmutz, J.; DiFazio, S. P.; Tschaplinski, T. J.; Tuskan, G. A.; Chen, J.-G.; Davison, B.; Ragauskas, A. J., Insights of biomass recalcitrance in natural Populus trichocarpa variants for biomass conversion. *Green Chemistry* **2017,** *19* (22), 5467-5478.

29. Heller, W. T.; Urban, V. S.; Lynn, G. W.; Weiss, K. L.; O'Neill, H. M.; Pingali, S. V.; Qian, S.; Littrell, K. C.; Melnichenko, Y. B.; Buchanan, M. V.; Selby, D. L.; Wignall, G. D.; Butler, P. D.; Myles, D. A., The Bio-SANS instrument at the High Flux Isotope Reactor of Oak Ridge National Laboratory. *Journal of Applied Crystallography* **2014,** *47* (4), 1238-1246.

30. Langan, P.; Petridis, L.; O'Neill, H. M.; Pingali, S. V.; Foston, M.; Nishiyama, Y.; Schulz, R.; Lindner, B.; Hanson, B. L.; Harton, S.; Heller, W. T.; Urban, V.; Evans, B. R.; Gnanakaran, S.; Ragauskas, A. J.; Smith, J. C.; Davison, B. H., Common processes drive the thermochemical pretreatment of lignocellulosic biomass. *Green Chemistry* **2014,** *16* (1), 63-68.

31. Pingali, S. V.; Urban, V. S.; Heller, W. T.; McGaughey, J.; O'Neill, H.; Foston, M. B.; Li, H.; Wyman, C. E.; Myles, D. A.; Langan, P.; Ragauskas, A.; Davison, B.; Evans, B. R., Understanding Multiscale Structural Changes During Dilute Acid Pretreatment of Switchgrass and Poplar. *ACS Sustainable Chemistry & Engineering* **2017,** *5* (1), 426-435.

32. Beaucage, G., Approximations Leading to a Unified Exponential/Power-Law Approach to Small-Angle Scattering. *Journal of Applied Crystallography* **1995,** *28* (6), 717-728.

33. Pingali, S. V.; Urban, V. S.; Heller, W. T.; McGaughey, J.; O'Neill, H.; Foston, M.; Myles, D. A.; Ragauskas, A.; Evans, B. R., Breakdown of Cell Wall Nanostructure in Dilute Acid Pretreated Biomass. *Biomacromolecules* **2010,** *11* (9), 2329-2335.

34. Phyo, P.; Wang, T.; Kiemle, S. N.; O'Neill, H.; Pingali, S. V.; Hong, M.; Cosgrove, D. J., Gradients in Wall Mechanics and Polysaccharides along Growing Inflorescence Stems. *Plant Physiology* **2017,** *175* (4), 1593-1607.

35. Shah, R.; Huang, S.; Pingali, S. V.; Sawada, D.; Pu, Y.; Rodriguez, M.; Ragauskas, A. J.; Kim, S. H.; Evans, B. R.; Davison, B. H.; O'Neill, H., Hemicellulose–Cellulose Composites Reveal Differences in Cellulose Organization after Dilute Acid Pretreatment. *Biomacromolecules* **2019,** *20* (2), 893-903.

36. Nishiyama, Y., Structure and properties of the cellulose microfibril. *Journal of Wood Science* **2009,** *55* (4), 241-249.

37. Fernandes, A. N.; Thomas, L. H.; Altaner, C. M.; Callow, P.; Forsyth, V. T.; Apperley, D. C.; Kennedy, C. J.; Jarvis, M. C., Nanostructure of cellulose microfibrils in spruce wood. *Proceedings of the National Academy of Sciences* **2011,** *108* (47), E1195-E1203.

38. Ilavsky, J.; Jemian, P. R., Irena: tool suite for modeling and analysis of small-angle scattering. *Journal of Applied Crystallography* **2009,** *42* (2), 347-353.

39. Pingali, S. V.; O'Neill, H.; Nishiyama, Y.; He, L.; Melnichenko, Y. B.; Urban, V. S.; Petridis, L.; Davison, B.; Langan, P., Morphological changes in the cellulose and lignin components of biomass occur at different stages during steam pretreatment. *Cellulose* **2014,** (2), 873–878.

40. Heller, W. T.; Urban, V. S.; Lynn, G. W.; Weiss, K. L.; O'Neill, H. M.; Pingali, S. V.; Qian, S.; Littrell, K. C.; Melnichenko, Y. B.; Buchanan, M. V., The Bio-SANS instrument at the high flux isotope reactor of Oak Ridge National Laboratory. *Journal of Applied Crystallography* **2014,** *47* (4), 1238-1246.

41. Heller, W. T.; Cuneo, M.; Debeer-Schmitt, L.; Do, C.; He, L.; Heroux, L.; Littrell, K.; Pingali, S. V.; Qian, S.; Stanley, C.; Urban, V. S.; Wu, B.; Bras, W., The suite of small-angle neutron scattering instruments at Oak Ridge National LaboratoryThis article will form part of a virtual special issue on advanced neutron scattering instrumentation, marking the 50th anniversary of the journal. *Journal of Applied Crystallography* **2018,** *51* (2), 242-248.

42. Driemeier, C.; Mendes, F. M.; Santucci, B. S.; Pimenta, M. T., Cellulose co-crystallization and related phenomena occurring in hydrothermal treatment of sugarcane bagasse. *Cellulose* **2015,** *22* (4), 2183-2195.

43. Hermans, P. H.; Heikens, D.; Weidinger, A., A quantitative investigation on the X-ray small angle scattering of cellulose fibers. Part II. The scattering power of various cellulose fibers. *Journal of Polymer Science* **1959,** *35* (128), 145-165.

44. Wang, H.; Keum, J. K.; Hiltner, A.; Baer, E., Impact of Nanoscale Confinement on Crystal Orientation of Poly(ethylene oxide). *Macromolecular Rapid Communications* **2010,** *31* (4), 356-361.

45. Fernandes, A. N.; Thomas, L. H.; Altaner, C. M.; Callow, P.; Forsyth, V. T.; Apperley, D. C.; Kennedy, C. J.; Jarvis, M. C., Nanostructure of cellulose microfibrils in spruce wood. *Proceedings of the National Academy of Sciences* **2011,** *108* (47), E1195.

46. Sun, Q.; Foston, M.; Meng, X.; Sawada, D.; Pingali, S. V.; O'Neill, H. M.; Li, H.; Wyman, C. E.; Langan, P.; Ragauskas, A. J.; Kumar, R., Effect of lignin content on changes occurring in poplar cellulose ultrastructure during dilute acid pretreatment. *Biotechnology for Biofuels* **2014,** *7* (1), 150.

47. Foston, M.; Ragauskas, A. J., Changes in lignocellulosic supramolecular and ultrastructure during dilute acid pretreatment of Populus and switchgrass. *Biomass and Bioenergy* **2010,** *34* (12), 1885-1895.

48. Sannigrahi, P.; Ragauskas, A. J.; Miller, S. J., Effects of Two-Stage Dilute Acid Pretreatment on the Structure and Composition of Lignin and Cellulose in Loblolly Pine. *BioEnergy Research* **2008,** *1* (3), 205-214.

49. Sun, Q.; Khunsupat, R.; Akato, K.; Tao, J.; Labbé, N.; Gallego, N. C.; Bozell, J. J.; Rials, T. G.; Tuskan, G. A.; Tschaplinski, T. J.; Naskar, A. K.; Pu, Y.; Ragauskas, A. J., A study of poplar organosolv lignin after melt rheology treatment as carbon fiber precursors. *Green Chemistry* **2016,** *18* (18), 5015-5024.

50. Chandra, R.; Ewanick, S.; Hsieh, C.; Saddler, J. N., The characterization of pretreated lignocellulosic substrates prior to enzymatic hydrolysis, part 1: A modified Simons' staining technique. *Biotechnology Progress* **2008,** *24* (5), 1178-1185.

51. Kuribayashi, T.; Ogawa, Y.; Rochas, C.; Matsumoto, Y.; Heux, L.; Nishiyama, Y., Hydrothermal Transformation of Wood Cellulose Crystals into Pseudo-Orthorhombic Structure by Cocrystallization. *ACS Macro Letters* **2016,** *5* (6), 730-734.

52. Kurašin, M.; Väljamäe, P., Processivity of Cellobiohydrolases Is Limited by the Substrate. *Journal of Biological Chemistry* **2011,** *286* (1), 169-177.

53. Igarashi, K.; Koivula, A.; Wada, M.; Kimura, S.; Penttilä, M.; Samejima, M., High Speed Atomic Force Microscopy Visualizes Processive Movement of Trichoderma reesei Cellobiohydrolase I on Crystalline Cellulose. *Journal of Biological Chemistry* **2009,** *284* (52), 36186-36190.

54. Igarashi, K.; Uchihashi, T.; Koivula, A.; Wada, M.; Kimura, S.; Okamoto, T.; Penttilä, M.; Ando, T.; Samejima, M., Traffic Jams Reduce Hydrolytic Efficiency of Cellulase on Cellulose Surface. *Science* **2011,** *333* (6047), 1279-1282.

55. Muchero, W.; Ehlers, J. D.; Close, T. J.; Roberts, P. A., Genic SNP markers and legume synteny reveal candidate genes underlying QTL for Macrophomina phaseolina resistance and maturity in cowpea [Vigna unguiculata (L) Walp.]. *BMC Genomics* **2011,** *12* (1), 8.

56. Foston, M.; Hubbell, C. A.; Davis, M.; Ragauskas, A. J., Variations in cellulosic ultrastructure of poplar. *BioEnergy Research* **2009,** *2* (4), 193.

57. Zhao, J. K.; Gao, C. Y.; Liu, D., The extended Q-range small-angle neutron scattering diffractometer at the SNS. *Journal of Applied Crystallography* **2010,** *43* (5 Part 1), 1068-1077.

58. Arnold, O.; Bilheux, J. C.; Borreguero, J. M.; Buts, A.; Campbell, S. I.; Chapon, L.; Doucet, M.; Draper, N.; Ferraz Leal, R.; Gigg, M. A.; Lynch, V. E.; Markvardsen, A.; Mikkelson, D. J.; Mikkelson, R. L.; Miller, R.; Palmen, K.; Parker, P.; Passos, G.; Perring, T. G.; Peterson, P. F.; Ren, S.; Reuter, M. A.; Savici, A. T.; Taylor, J. W.; Taylor, R. J.; Tolchenov, R.; Zhou, W.;

Zikovsky, J., Mantid—Data analysis and visualization package for neutron scattering and μ SR experiments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2014,** *764*, 156-166.
59. Cardoso, M. B.; Luckarift, H. R.; Urban, V. S.; O'Neill, H.; Johnson, G. R., Protein Localization in Silica Nanospheres Derived via Biomimetic Mineralization. *Advanced Functional Materials* **2010,** *20* (18), 3031-3038.
60. Nishiyama, Y.; Langan, P.; Chanzy, H., Crystal Structure and Hydrogen-Bonding System in Cellulose Iβ from Synchrotron X-ray and Neutron Fiber Diffraction. *Journal of the American Chemical Society* **2002,** *124* (31), 9074-9082.
61. Li, Y.; Fu, Q.; Yu, S.; Yan, M.; Berglund, L., Optically Transparent Wood from a Nanoporous Cellulosic Template: Combining Functional and Structural Performance. *Biomacromolecules* **2016,** *17* (4), 1358-1364.

Chapter 6). A study to probe associations between lignin and homogalacturonan

6.1). Introduction

Biomass recalcitrance to enzymatic hydrolysis is often attributed to lignin, a complex phenylpropanoid polymer making about 20-25% of the plant cell wall.[1] At the industrial scale, the rigid structure of the plant cell wall is made less recalcitrant by pretreating it with solvents such as dilute acid or water at elevated temperatures.[2] Several studies have shown that after such pretreatments, lignin forms aggregates and cellulose degrading enzymes bind unproductively to these aggregates.[3] The lignin aggregates prevent the enzymes from accessing cellulose surface leading to incomplete digestion of biomass.[4] Different interpretations explaining the mechanism through which the lignin collapses and forms aggregates. Some studies attribute these changes in morphology of lignin post pretreatment to the increase in the degree of lignin condensation and delocalization[5] while some interpret it as lignin melting, migration and re-deposition as lignin droplets.[6] The association of polysaccharides with lignin and the role polysaccharide play in changing lignin morphology during pretreatment has also been questioned.[7] For instance, DeMartini et al used monoclonal antibodies to perform glycome profiling of poplar cell wall at post different stages of hot water pretreatment and showed that as different polysaccharides get released during pretreatment, it led to lignin rearrangement. [7] Their study was suggestive that lignin polysaccharide interactions play a role in recalcitrance.

Lignin-polysaccharide interactions were first proposed about 150 years ago by Erdman when he found that there was a fraction of lignin and carbohydrates that were inseparable.[8] Such associations are referred to in the literature as lignin-carbohydrate complexes. Several research groups have tried to isolate the lignin carbohydrate complexes from the plant cell wall in order to know the type of polysaccharide residue involved in making such linkages.[9-14] However, as harsh chemicals were used in the extraction process, it was argued that the LCCs may have formed as a result of use of these chemicals and might not be present in the native plant cell wall.[13] Several milder processes to isolate the lignin carbohydrate complexes have been developed but the LCCs isolated exist in such small amounts that analyzing them has been difficult.[15] The whole plant cell wall has been studied with solid state NMR in order to identify the LCCs but due to a large number of cross peaks in the NMR pattern it was difficult to assign the peaks.[9] According to the current view of polymer interactions in the plant cell wall, it is thought that lignin interacts with cellulose through the hemicellulose and pectin network.[16] Therefore, interactions of lignin with polysaccharides like pectin and hemicellulose are thought to be participating in forming LCCs.

At various stages of the development of the plant cell wall, different polymers get formed and assembled into the plant cell wall.[17] Lignin formation starts after the primary cell wall has formed. While lignin formation starts in the middle lamella and at the cell corners, it is found in abundance in the secondary cell wall of the plants along with hemicellulose. Ester linkages between arabinoxylan, a type of hemicellulose and ferulic acid containing lignin have been found in grasses.[18-19] Recently, using solid state NMR on whole plant cell walls, it was shown that hemicellulose interacts with lignin through electrostatic interactions.[20] Pectin’s binding properties to cellulose have been studied and inferences about its role in cell wall accessibility[21] have been made but pectin’s interactions with lignin are largely unexplored. One reason for this could be that while pectin is a major primary wall component, its presence in the secondary cell wall is almost negligible.[22] Despite pectin having unique sugar monomers like rhamnose, galacturonic acid, glucuronic acid, apiose, its content is so low that it is not even reported in the sugar compositional analysis of woody cell walls used for making cellulosic biofuel. Although the pectin content is low, for the isolated LCCs from softwood it was found that they had galactose and arabinose content which was relatively higher than the average hemicellulose composition of softwood. β-D-1->4 galactan and α-l-1->5-arabinan are characteristic groups of pectin and the isolated LCC fraction have suggested to be pectin-lignin bond.[23]

A recent study showed that downregulation of the galacturonosyltransferase 4 (GAUT 4) enzyme which is needed for the α- 1,4 links between galacturonan to form homogalacturonan led to increased switchgrass biomass and glucose yield (~24 to 38%) post hydrolysis.[24] The study also used specific antibodies to show that homogalacturonan mainly exists in the middle lamella and at plant cell wall corners (Figure 6.1B). Interestingly, when lignin is removed from wood by delignification a hollow lumen is seen[25] in the exact area where homogalacturonan is shown to exist (Figure 6.1C). This indicates that lignin and homogalacturonan are co-located in the middle lamella region of the plant cell wall.

Additionally, DeMartini et al. (mentioned above),[7] conducted a glycome profiling study to check for the presence of different polysaccharides remaining in the cell wall post different severity of hot water pretreatment. They found that for the mild hot water pretreatment conditions, one of the first polysaccharides to be removed from cell wall was arabinogalactan and homogalacturonan.[7] They also report that upon removal of these polysaccharides, the lignin was re-distributed within the plant cell wall. The co-location of lignin and homogalacturonan along with mild pretreatment conditions leading to the removal of homogalacturonan and causing lignin rearrangement made us interested in exploring existence of lignin carbohydrate complexes between lignin and homogalacturonan.

We have previously observed the changes in lignin distribution due to hot water pretreatment of switchgrass and poplar using small-angle neutron scattering (SANS).

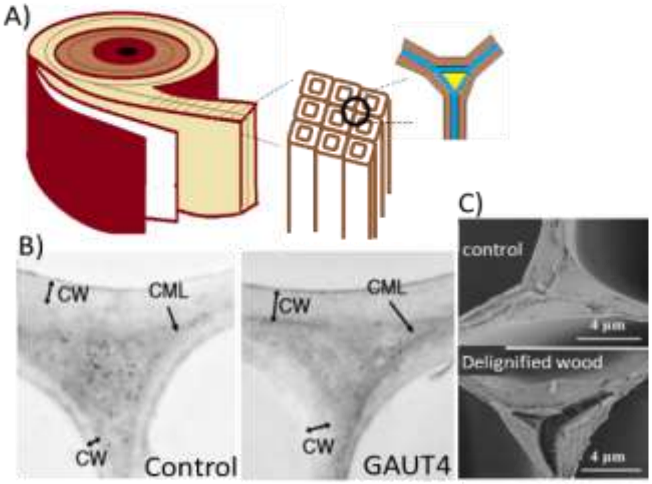


Figure 6.1: A). cartoon representation of the middle lamella region B). black spots obtained by using antibodies that specifically bind homogalacturonan in the middle lamella (CML) region[7] C). Lignified and delignified wood[25]

In order to see if lignin would re-distribute differently in wildtype (WT) and GAUT4- knockdown (kd) switchgrass after hot water pretreatment, we collected the SANS profile. Our hypothesis was if that the presence of homogalacturonan can delay the formation of lignin aggregates as compared to when it is not present then an association between homogalacturonan and lignin can be shown to exist. The rationale was that in order for lignin to re-distribute and form aggregates post hot water pretreatment the interactions between lignin and homogalacturonan must be disrupted. We also polymerized monolignols in presence of homogalacturonan for providing supporting evidence that they could be linked by covalent linkages and used spectroscopy to determine the type of linkage between them.

6.2). Materials and method

6.2.1). Reagents for dehydrogenation polymer

Horseradish peroxidase (specific activity ≥ 250 units/mg solid, type IV, P8375-1KU), hydrogen peroxide solution (30% wt in H_2O), coniferyl alcohol (molecular weight 180.20 CAS number:458-35-5) were purchased from Millipore-Sigma.

6.2.1.1). Pectin

Polygalacturonic acid (Na-Salt, CAS number: 9049-370) extracted from citrus peel was purchased from Megazyme (Ireland). 1% (w/v). Pectin solutions were prepared in water or in 50 mM sodium acetate solution (pH 5.0) and were stirred overnight to ensure complete dissolution.

6.2.1.2). Lignin dehydrogenation polymer (DHP) synthesis

The synthesis of DHP was done by following the previously published procedure with a slight modification. Coniferyl alcohol solution (34mM) was prepared by dissolving 200 mg of coniferyl alcohol in 2 mL of dioxane and the volume was adjusted to 20 mL by addition of H_2O. Hydrogen peroxide (H_2O_2) solution was prepared by adding 175 ul H_2O_2 (30% w/v) to 20 mL H_2O. 1 mg of horse radish peroxide (HRP) was dissolved in 100 mL of 50 mM sodium acetate buffer (pH 5.0). This solution was constantly stirred at 400 rpm and the coniferyl alcohol solution and H_2O_2 solution were added to it. Both, coniferyl alcohol solution and H_2O_2 solution were added dropwise using a syringe pump at flow rate of 5-7 mL/h. Once the addition was complete, the mixture of the solutions was stirred for 24 h and was then centrifuged at 12,000 rpm for 30 mins. The pellet and the supernatant were separated and the pellet was washed three times were 20-30 mL water in order to ensure removal of dioxane. The pellet was then taken and freeze-dried overnight.

6.2.2). DHP-pectin synthesis

The DHP-pectin was prepared by dissolving pectin (1 gm) to 100 mL of 50 mM of sodium acetate buffer or 100 mL of water and was stirred for 24 hrs to ensure complete dissolution of pectin. To this solution 1 mg of HRP was added and the procedure was synthesizing DHP was followed. The mixture of the reactants and formed products were centrifuged at 12,00 rpm for 30 mins. The supernatant and the pellet were separated. The pellet was washed several times and both the supernatant and pellet were freeze-dried

6.2.3). Acetylation of DHP

50 mg of the freeze-dried DHP sample was transferred into a glass vial. Acetic anhydride (1 mL) and pyridine (1mL) were added in equal ratio and stirred over-night. The reaction was then transferred in oven dried flask and 30 ml anhydrous ethanol was added. The reaction mixture and ethanol were evaporated using a rotary evaporator. To ensure complete removal of acetic anhydride and pyridine, the evaporation step with ethanol was repeated three times.

6.2.4). Molecular weight determination of DHP

The molecular weight was determined using the high-performance liquid chromatography (Hewlet-Packard 1090 series system) coupled with UV detector. The acetylated DHP was loaded onto three Styragel columns (HR1, HR3, and HR4 from Waters) which were linked in series and the elution was done using tetrahydrofuran. The column was calibrated using pullulan standards of molecular weight 708 kDa, 344 kDa, 194kDa, 47.1 kDa, 21.1 kDa, 9.6 kDa, 6 kDa, 1.08 kDa, and 342 kDa.

6.2.5). Infrared spectroscopy of DHP and DHP-pectin

The bonds in the freeze-dried pectin, DHP and pectin-DHP composites were determined using attenuated reflectance transmission (ATR) infrared (IR) spectrophotometry (Jasco, 6100). The IR spectra were obtained by scanning wavenumber from 4000 cm^{-1} to 500 cm^{-1}. The sample was placed on the zirconium crystal and pressed on the crystal surface, and three different regions in the sample were measured. Each measurement was an average of 16 scans and resolution was 4 cm^{-1}. Corrections, such as baseline and, removal of carbon-dioxide interference, were done using the Jasco spectra manager (version 2.0).

6.2.6). Small angle x-ray scattering of pectins and DHP-pectin

The X-ray scattering measurements were performed on an in house SAXS instrument (Bio-SAXS-2000 Nanostar, Rigaku). The X-rays were being generated by a rotating anode by operating the instrument at a voltage of 40 kV and a current of 30 mA. A Kratky block was used to collimate the CuKα X-ray beam of wavelength 1.54 Å and achieve a Q-range of 0.006 to 0.6 $Å^{-1}$. The beam size at the sample was 1.5 mm^2 and the sample to detector distance was 480 mm. This was calibrated using sodium behenate as a standard. The instrument had a pixel array detector (Riguku

HyPix-3000) along with a temperature-controlled sample stage for glass capillaries. The pectins and DHP-pectins samples were loaded into a 1mm thick glass capillary and measured for 1 hr. The process of conversion of measured data 2D data to a 1D SAXS profile was an in-built feature of the data processing software provided with the instrument. The scattering from solvent in the capillary was used as background which was subtracted from the sample scattering using the data processing software from Rigaku and IRENA. The data were fit with different models using IRENA.

6.2.7). Pretreatment and delignification of the GAUT4 and control switchgrass

For hot water pretreatment, ground native GAUT4 and control switchgrass were individually processed, by first soaking 2.12 gms of ground material overnight in 40 ml of distilled water. The presoaked slurry was transferred to a 75 mL total capacity Parr reactor with a 5% dry solid loading. The reactor was sealed and heated to 180 ˚C. The system took ~25 min to reach 180 ˚C and old at this temperature for 18min. The reaction was quenched by placing the sealed Parr reactor on ice until the temperature reached 25 ˚C. The pretreated wood slurry was recovered and washed thoroughly with distilled water. Delignification of native and pretreated wood samples was carried out as previously described.

6.2.8). Small angle neutron scattering of the GAUT4 and control switchgrass

The SANS data on delignified native and pretreated samples were collected at the EQ-SANS beamline located at the Spallation Neutron Source (SNS) which was operating at 60 Hz. The samples were packed with the ground GAUT4 and control switchgrass into a titanium cell with thickness of 1mm and filled with 100% D_2O. The neutron beam was collimated using an aperture size of 10 mm and the scattering data was collected by using sample of detector distance (SDD) of 2.5 m and 4m. The neutron wavelength used for the two SDDs were 2.5 and 10 Å respectively. The combination of the SDD and neutron wavelength enabled us to obtain a Q range spanning 0.005 $Å^{-1}$ < Q <$0.7Å^{-1}$. The data was reduced by using the standard procedures implemented in the Mantid software[46] and an output file containing Q, I(Q) and error in Q and I(Q) was obtained. The intensity I (Q) in the output file was calibrated to be on absolute scale using porous silica as a standard.

The SANS data were fit to the multi-level Unified Fit model implemented in the IRENA package of Igor Pro software by Wavemetrics. As shown in Equation 1 the SANS intensity profile is a summation of individual levels (i) with each level modeled as the sum of an exponential and a power law behavior

$$I(q) = \sum_i I_i(q) = G_i e^{-\frac{q^2 R_{g_i}^2}{3}} + e^{-\frac{q^2 R_{gCO_{i-1}}^2}{3}} \times B_i \left\{ \frac{\left(\mathrm{erf}\left(\frac{qR_{g_i}}{\sqrt{6}} \right) \right)^3}{q} \right\}^{P_i} + C$$

where R_{g_i} is the radius of gyration of the particle of the i^{th} structural level, P_i is the power-law exponent of the i^{th} structural level, G_i is the scalar for the Guinier function of the i^{th} structural level, and B_i is the scalar for the power-law function of the i^{th} structural level. Additionally, $R_{gCO_{i-1}}$ is the cut-off length scale of the power-law behavior of the i^{th} structural level and C is the q-independent constant background intensity.

6.3). Results

6.3.1). Small angle neutron scattering (SANS) analysis of native and hot water pretreated wildtype and GAUT4-kd Switchgrass

SANS was used to study the morphological changes in the cell wall structure of native and liquid hot water pretreated wildtype (WT) and GAUT4-kd switchgrass. The 1D SANS profile of native wildtype and GAUT4-kd switchgrass are shown in Figure 6.2 panel A.

For the Q-range measured, these curves show two distinct structural regimes namely the primary regime (Level-1) from 0.08 $Å^{-1} < Q < 0.4$ $Å^{-1}$and secondary regime (Level-2) from 0.0056 $Å^{-1} < Q < 0.08$ $Å^{-1}$. The unified model gives the ability to sum different structural levels and obtain a fit for measured Q range of 0.0056 $Å^{-1} < Q < 0.4$ $Å^{-1}$. The solid black lines in Figure 6.2 is the overall fit for the scattering curves. Each regime was fit by using a radius of gyration (Rg) and power law exponent (P) and the parameter values are summarized in Table 6.1.

The primary level scattering is mainly due to cellulose microfibril arrangement while the secondary level scattering is due to the cell wall polymer network. As seen from Figure 6.2 panel A, no difference between the scattering curves are seen between WT and GAUT4-kd switchgrass. A similar cross-sectional size (level 1-Rg) was determined for the WT and GAUT4-kd switchgrass under native conditions. The power law exponent (2P) was suggestive of a similar type of highly branched network structure in both these switchgrass samples.

The native WT and GAUT4-kd switchgrass were subjected to liquid hot water pretreatment (LHW) in the Parr reactor as described in the materials and method part. The scattering curves of the pretreated wildtype and GAUT4-kd switchgrass are shown in Figure 6.2 panel B. The Q-region from ~0.015 to 0.08 no longer follows the power law scattering rather a particle size feature was visible in these samples post pretreatment.

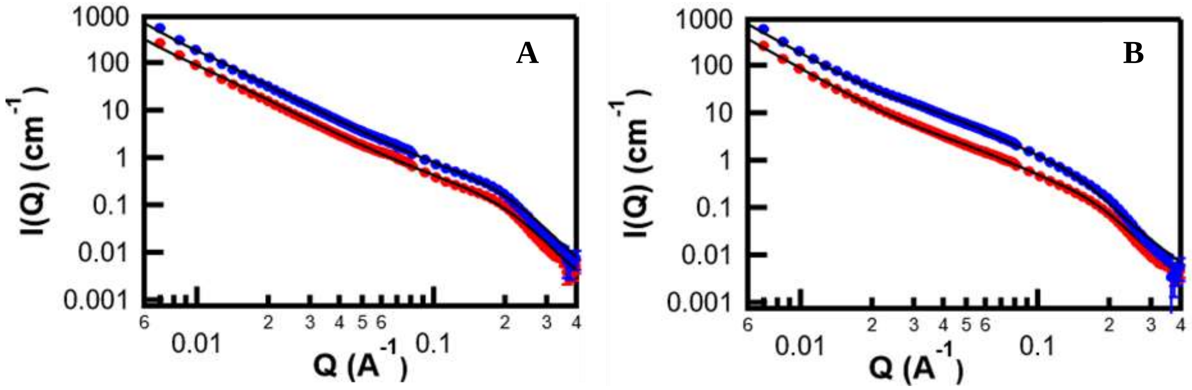


Figure 6.2: Scattering curves and fit (black) for the native (left) and hot water pretreated (right) wildtype (red) and GAUT4 (blue) switchgrass

Table 6.1: Fitting parameters obtained from SANS of Wildtype and GAUT4-kd switchgrass in 100% D$_2$O

	Level 1		**Level 2***	
Switchgrass	1G	1Rg (Å)	2B	2P
Wildtype	0.358	11.0+/- 1.2	0.00093	2.45 +/- 0.06
GAUT4-kd	0.2814	11.0 +/- 0.7	0.000701	2.52 +/- 0.03
Level 1 = 0.08 Å^{-1}< Q < 0.4 Å^{-1}; Level 2 = 0.0056 Å^{-1}< Q < 0.08 Å^{-1} Error bar obtained by uncertainty analysis; *2Rg (500) and 2G (2000) fixed				

Due to the change in scattering features between the native and pretreated switchgrass samples, three structural levels were needed to fit the scattering Q-range with unified fit. The primary regime was from 0.08 $Å^{-1} < Q < 0.4$ $Å^{-1}$ (Level-1), secondary regime from 0.02 $Å^{-1} < Q < 0.08$ $Å^{-1}$ (Level-2) and the tertiary regime was from 0.0056 $Å^{-1} < Q < 0.02$ $Å^{-1}$(Level-3). The fit results are summarized in Table 6.2.

The cellulose microfibril cross-sectional section size has slightly increased (Level-1Rg) to 14.34 +/- 0.45 Å after hot water pretreatment in the wildtype switchgrass than in its native state (Rg = 11.2 +/- 1.2 Å). This indicates the coalescing of cellulose microfibrils and is consistent with previous publications. A similar increase was also seen for the hot water pretreated GAUT4-kd switchgrass with a Rg value was 16.30 +/- 0.73 Å. A particle feature was observed in the secondary regime (0.025 $Å^{-1} < Q < 0.08$ $Å^{-1}$) and was fit to using a Guinier function. It was previously shown that post delignification of hot water pretreated switchgrass, the feature was no longer present in the scattering profile and due to this it was attributed to lignin aggregates.

The lignin aggregate size for a hot water pretreated wildtype was 35.1 +/- 1.9 Å and for the GAUT4-kd was 39.1 +/- 2.4 Å. While a slight difference is seen in the size of the aggregate, the scattering curve of pretreated GAUT4-kd has a more pronounced intensity than pretreated wildtype. The intensity indicates the number of lignin aggregates and was quantified using a scalar (G). There was a two-fold fewer lignin aggregates in the pretreated wildtype (G~3) compared to pretreated GAUT4-kd switchgrass (G~6).

Table 6.2: Fitting parameters obtained from SANS of hot water pretreated wildtype and GAUT4-kd switchgrass in 100% D_2O

	Level 1		**Level 2**		**Level 3**	
Switchgrass	1G	1Rg (Å)	2G	2Rg (Å)	3B	3P
LHW Wildtype	0.78	14.3 +/- 0.5	2.75	35.0 +/- 2	**1.98E-04**	**2.82**
LHW GAUT4-kd	1.44	16.3 +/- 0.7	6	39.0 +/- 2	**0.0002104**	**2.807**
* Bold = parameters fixed; Error bar obtained by uncertainty analysis Level 1 = 0.08 $\text{Å}^{-1} < Q < 0.4\ \text{Å}^{-1}$; Level 2 = 0.02 $\text{Å}^{-1} < Q < 0.08\ \text{Å}^{-1}$; Level 3 = 0.0056 $\text{Å}^{-1} < Q < 0.02\ \text{Å}^{-1}$						

The lignin-like polymer was made by slow addition of coniferyl alcohol to a solution containing horseradish peroxidase. As it was not water-soluble and it was obtained by centrifugation of the solution. The lignin-like polymer product that is formed by dehydrogenation of coniferyl alcohol is called dehydrogenation polymer (DHP). A schematic explaining the process of the formation of DHP is shown in Figure 6.3.

Figure 6.4 shows an infrared spectrum of DHP and Klason lignin obtained from softwood. The DHP transmission IR spectra for DHP is quite similar to Klason lignin and different from its monomer, coniferyl alcohol. The DHP had ~25-30% aryl ether (β-O-4) linkages along with phenyl coumarin (β5) and pinoresinol (ββ) linkages as determined by NMR spectroscopy. The molecular weight of DHP determined by gel permeation chromatography (GPC) was ~1000 Dalton with a polydispersity of 1.5 indicating the DHP was made of roughly 5 to 6 monomers of coniferyl alcohol.

6.3.3). Dehydrogenation polymer-homogalacturonan (DHP-HG) composite

DHP-HG composites were prepared by polymerizing monolignol in a solution containing homogalacturonan and horseradish peroxidase. As DHP was not water-soluble, most of it was precipitated out of the DHP-HG solution. The solution was centrifuged and the supernatant was collected to ensure removal of all loosely bound DHP. The supernatant was used for IR study.

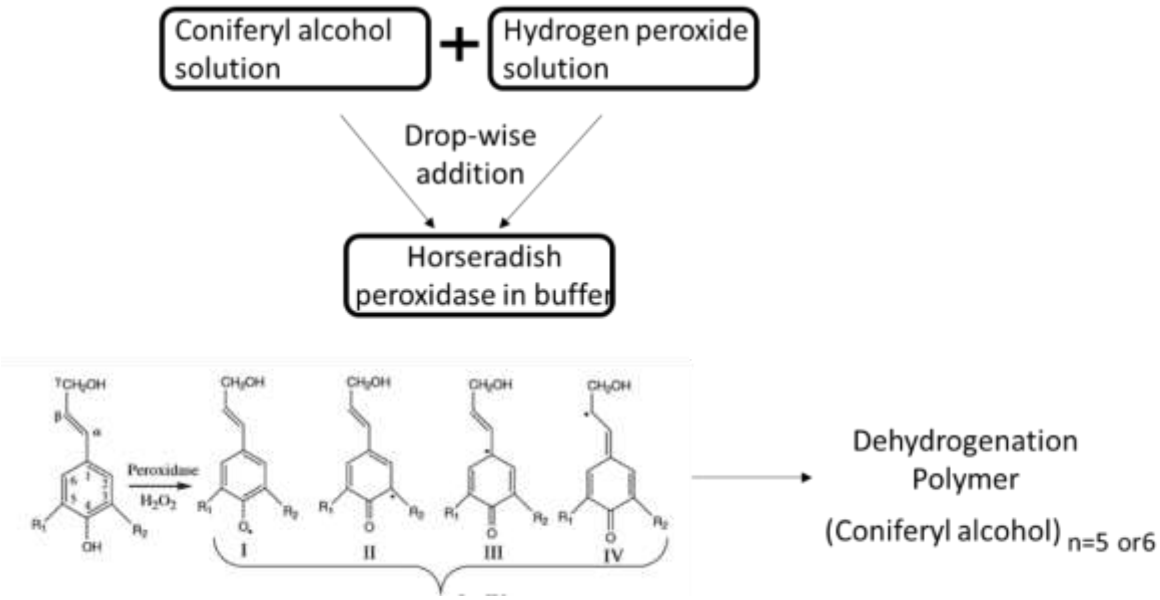


Figure 6.3: Horse radish peroxidase converts the coniferyl alcohol to free radical intermediate in presence of hydrogen peroxide that combine to from 5 to 6 mer long polymer called dehydrogenation polymer

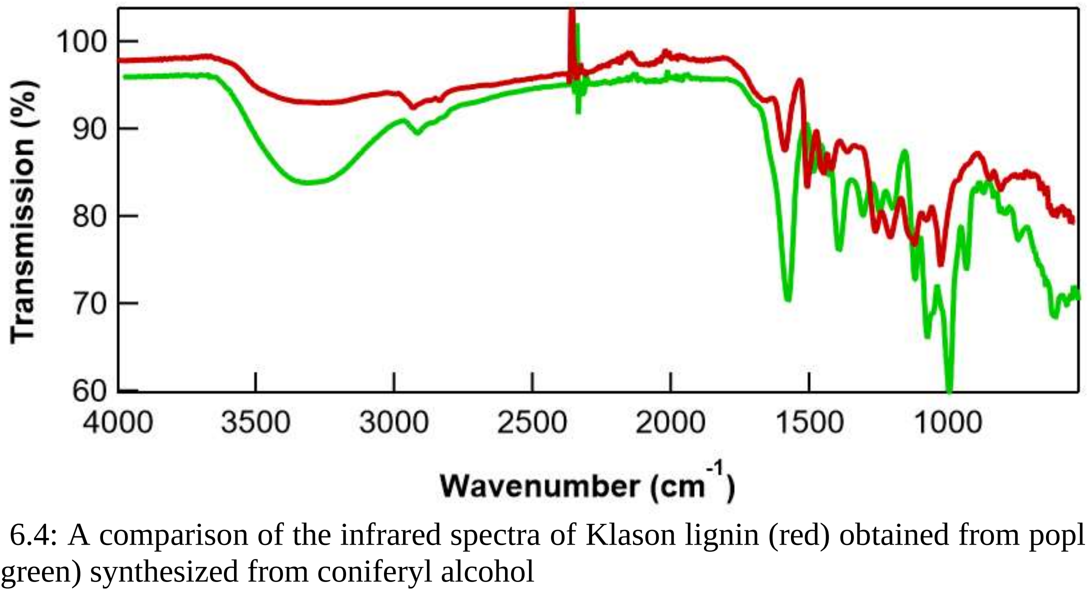


Figure 6.4: A comparison of the infrared spectra of Klason lignin (red) obtained from poplar and DHP (green) synthesized from coniferyl alcohol

Figure 6.5 shows IR spectra of a mixture of HG and DHP (HG-DHP mix) and DHP synthesized in the presence of HG (DHP- HG composite) along with HG only and DHP only. The spectra for DHP-HG mix looks similar to the spectra of HG only. The DHP- HG composite spectra look like a summation of HG only and DHP only spectra. The presence of DHP bands in the DHP-HG composite indicates that while most of the DHP was in the insoluble fraction, the supernatant fraction of DHP-HG composite still has DHP. Additionally, a unique band appears in the DHP-HG composite spectra at ~1730 cm^{-1}. A band was previously observed (Figure 6.6) in this region and occurs due to carbonyl stretching of the ester bond [26-27] suggestive of an ester linkage between DHP and HG.

6.3.4). Small angle x-ray scattering of different concentration of HG

The scattering curves obtained by measuring homogalacturonan at different concentrations are shown in Figure 6.7. A flat scattering in the low Q-range (0.01 to 0.06 Å^{-1}) of the HG solution at 1% (w/v) was observed. When the concentration was increased to 2% (w/v), an upturn in scattering is seen in the low Q-range (0.01 to 0.06 Å^{-1}). This type of scattering is typical for polymers forming polymer network in solution for which the low Q is due to cluster formation of polymer chain and the high Q (0.06 to 0.4 Å^{-1}) feature is due to polymer chain. The size of the cluster is beyond the measured Q-range and cannot be determined. Previously reported[28] functional form (eq 7.1) which was obtained by summing Porod scattering from clusters and Lorentzian function from polymer chain was used to determine the polymer network of 2% (w/v) HG chains in solution.

$$I(Q) = \frac{A}{Q^n} + \frac{C}{1 + (QL)^m} + B \qquad (7.1)$$

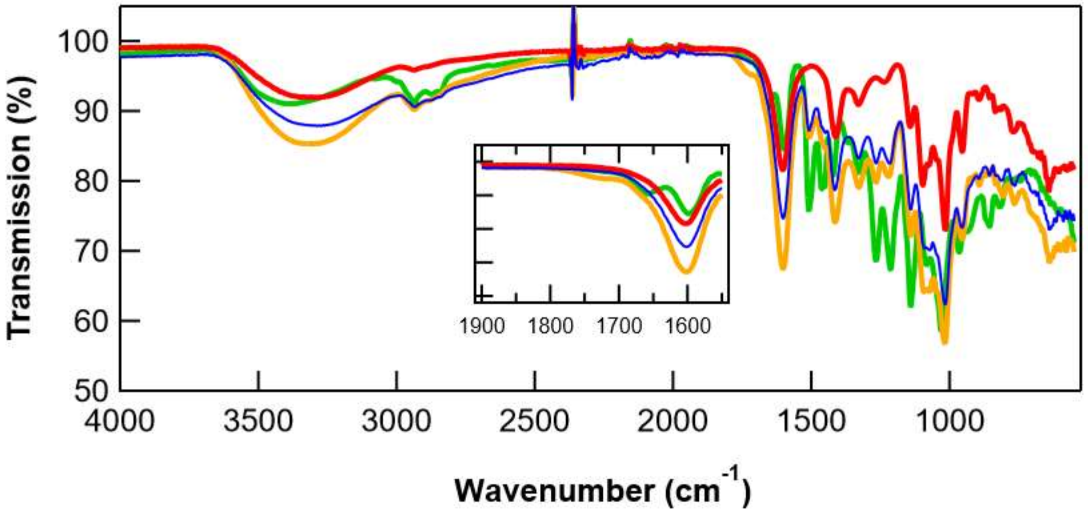


Figure 6.5: Infrared spectra of HG (red), synthesized DHP (green), DHP synthesized in presence of HG as HG-DHP composite (orange)and physical mixture of DHP and HG (blue).

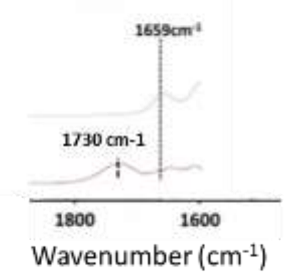


Figure 6.6: Previously reported ester linkage between DHP and cellulose nanocrystals (purple) at 1730 cm^{-1}. [26] The light blue curve is for DHP only

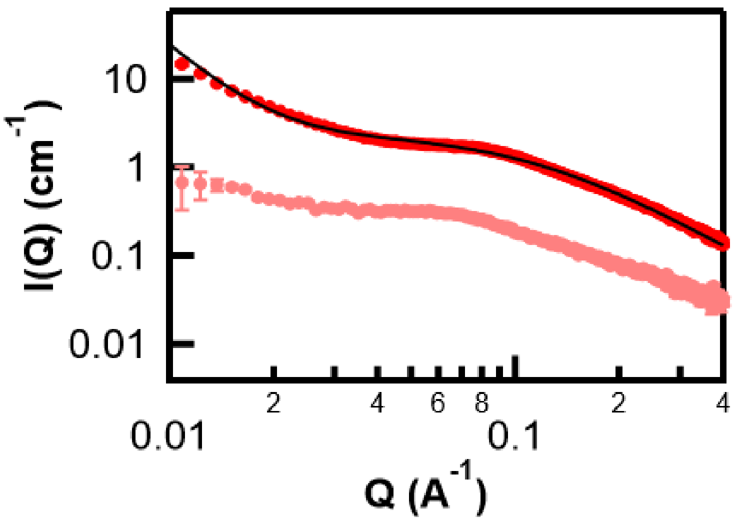


Figure 6.7: Small angle X-ray scattering from 1% (pink) and 2% (w/v) (red) homogalacturonan (HG) polymer in water at 25 ˚C. The solid black line was the fit obtained for 2% (w/v) HG to the experimental data using a model that is the sum of Porod scattering and Lorentzian function.

Here, L is the correlation length for the polymer chain, A and C are two scalar factors, B is the incoherent background and *n* and *m* are the exponents. The fit results obtained by using the functional form show the correlation (L) to be 8.86 +/- 0.04 Å and power law exponent (n) was 3.37+/- 0.03. The power law exponent of greater than 3 but less than 4 shows the surface fractal with rough interfaces with the solvent. Additionally, from the correlation length, the 2% HG solution appears to be a densely packed polymer network.

On the other 1% HG polymer does not form a dense network. The low Q porod scattering is flat indicative of individual polymers in solution. The cross-section size (Rc) of individual polymers can be derived by fitting the scattering data to the following equation (6.2).

$$QI(Q) \cong I(0)exp^{(-\frac{1}{2}Rc^2Q^2)} \tag{6.2}$$

Instead of the exponential form, the equation can be represented in the logarithmic form and Rc can be determined by representing the scattering curve (I(Q) vs Q) as a Guinier plot (ln (QI(Q)) vs Q^2)). The slope to a linear fit done to the scattering curve on such a plot is equal to the Rc. The Q range for which the equation can be used must satisfy the product of Q_{min} and Rc to be less than 1. The region is which this approximation holds true is shown in Figure 6.8.

It is also assumed that the polymer chains are rigid cylinders with radius (Rc) << length of the cylinder. The Rc value so determined was found to be 3.1 +/- 1 Å.

6.3.5). Small angle x-ray scattering of DHP and DHP-HG composites

The 1% HG solution was used to prepare the DHP-HG composite and the DHP-HG mix samples since at this concentration the HG polymer was individually suspended in solution.

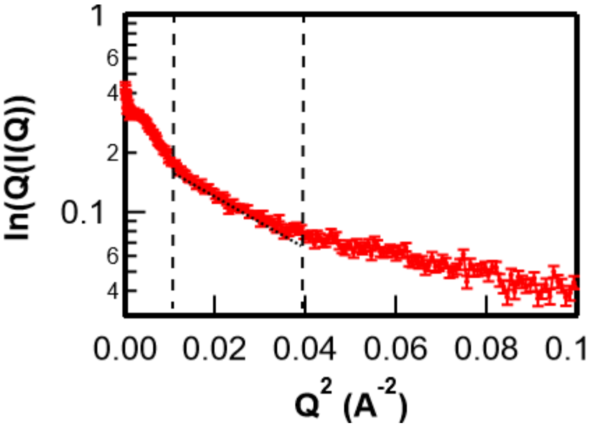


Figure 6.8: Guinier plot for the SAXS scattering profile of 1% (w/v) HG. The dashed vertical lines show the fitting range in which a linear fit (dashed fit line) was obtained.

The scattering curves for DHP-HG composite and DHP-HG mixture are shown in Figure 6.9. For comparison, the DHP only and HG only are also shown. For DHP-HG mixture, although 1% HG was used, there is an upturn in the low Q region. The DHP-HG mixture scattering curve visually looks like an addition of DHP only and HG only curve. The Unified fitting model was used to fit the DHP-HG mixture and DHP only curve (Figure 6.10). However, the upturn in low Q (0.01 $Å^{-1}$ < Q < 0.03 $Å^{-1}$) has an exponent value of 2.9 +/- 0.1 while the exponent value of DHP only has an exponent value of 4.5. Hence it is not conclusive to say whether DHP-HG mixture is like HG only with slightly higher concentration or it is an addition of scattering from DHP only and 1% HG only. Nonetheless, the scattering profile is very different from DHP-HG composite. The fitted scattering curve for DHP-HG composite is shown in Figure 6.10.

The Q-range of 0.01 to 0.3 $Å^{-1}$ of the curve was fit using two levels of the Unified fit. A particle size (Rg) of 35 +/- 9 Å along with a high Q powerlaw exponent of 2.95 +/- 0.56 and low Q powerlaw with exponent value 2.5+/-0.1 was determined from the fit. For the DHP-HG composite, there is an interaction of the hydrophobic DHP with the hydrophilic pectin. DHP remains in the HG solution and forms aggregates.

6.3.6). Analysis of ultra-small angle neutron scattering

Ultra-small angle neutron scattering was used to study the morphological changes in the cell wall structure of native and liquid hot water pretreated WT and GAUT4-kd switchgrass at microscopic scale. The measured total Q range from 0.00005 to 0.004 $Å^{-1}$ which in real space corresponds to 12.5 to 0.157 µm (using $d = \frac{2\pi}{Q}$). The 1D SANS profile of native wildtype and GAUT4-kd switchgrass before and after hot water pretreatment are shown in Figure 6.11.

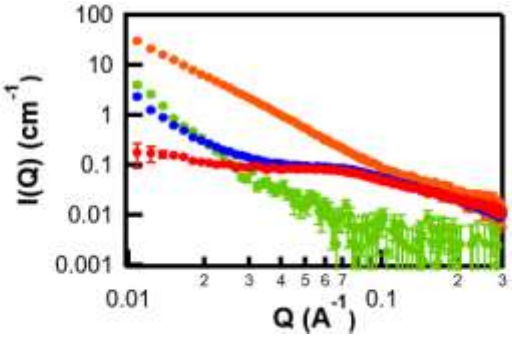


Figure 6.9: Difference in scattering pattern of HG (red), synthesized DHP (green), DHP synthesized in presence of HG (HG-DHP composite, in orange) and physical mixture of DHP and HG (DHP-HG mix, in blue). All the samples were measured in 100% H_2O

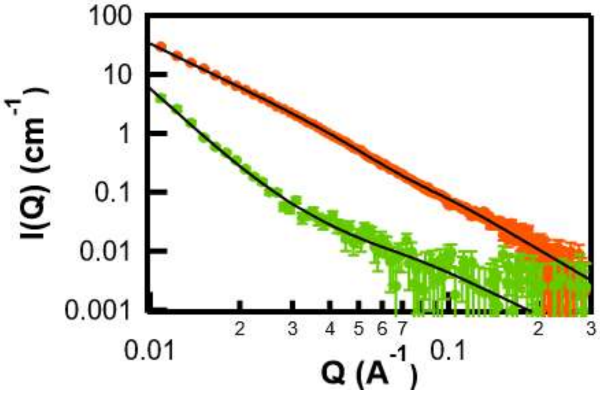


Figure 6.10: The solid black lines correspond to the fits obtained to the DHP only and DHP-HG composite scattering curves. Each curve was fit with 2 levels of Unified fit. The level 1 for DHP-HG and DHP only spans from Q-range of 0.003 to 0.3 Å^{-1} while level 2 spans from 0.003 to 0.01 Å^{-1}.

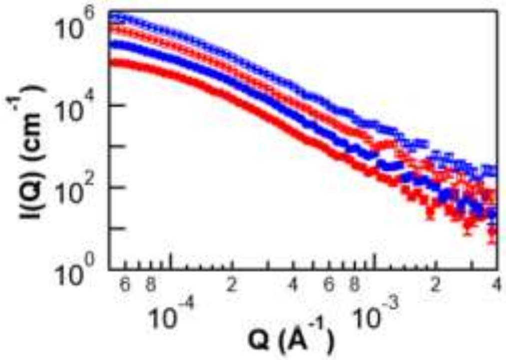


Figure 6.11: USANS curves for WT (closed red circle), GAUT4 (closed blue circle), LHW-WT (open red circle) and LHW-GAUT4 (open blue circle) switchgrass ground samples in 100%D_2O

Visually the scattering curves look very similar and fitting the data will allow us to conclusive determine if there are any changes. USANS of wood to study its microscopic structure was previously done for pine wood (P. radiata, a type of softwood) in which the scattering profile of wood was measuring from sections obtained by cutting the wood in transverse, orthogonal and parallel to the truck.[31] A similar approach will be taken to fit the scattering curve.

6.4). Discussion

In this study, three-year-old wildtype and galacturonosyl transferase 4 (GAUT4) knockdown switchgrass were studied with SANS. Structural information for the organization of cellulose microfibrils and the degree of crosslinking in the network formed by the cell wall polymers has been reported by studying switchgrass soaked in D_2O with SANS.[29] Our results obtained for both these structural features by fitting the wildtype switchgrass scattering curve are similar (within error) with the previous report. Our results obtained for both these structural features by fitting the wildtype switchgrass scattering curve are similar (within error) with the previous report. The scattering curve of the GAUT4-kd switchgrass is exactly the same as the wildtype switchgrass. This shows the lack of galacturonan containing pectin does not affect the cellulose cross-sectional size nor the network formed by the cell wall polymers. However, previously polysaccharides from the cell wall were sequentially extracted using chemical solutions in increasing degree of harshness. After each extraction step the cell wall was stained with different monoclonal antibodies to check for the presence of the polysaccharide that was remaining behind and indirectly determine the polysaccharide which got removed in the extraction process. It was much easier to extract certain polysaccharides from GAUT4-kd cell wall than from wildtype cell wall.[24] Due to this we expected to see the GAUT4-kd to have a loose polymer network than wildtype. This discrepancy in our expected and experimental result may be because the scattering profile is due to the contribution of all cell wall polymers. It could be that the polysaccharide network formed due to hemicellulose and pectin for GAUT4-kd is loose but the technique used is not sensitive enough to extract the contribution of these individual polymers in the network. Rather from the scattering profile, the polymer network formed by all cell wall polymers is visualized which is similar in both wildtype and GAUT4-kd.

To observe if a difference exists in the lignin aggregation pattern of wildtype and GAUT4-kd switchgrass, both these samples were subjected to hot water pretreatment for the same amount of time and temperature. Hot water pretreatment is known to agglomerate lignin and form lignin aggregates. They can clearly be visualized in the scattering profile as the midQ region in the scattering profile transitions from linear scattering to showing a particulate size. While the aggregate size in the post pretreated wildtype and GAUT4-kd switchgrass sample was similar, the number of such aggregates are more in GAUT4 than they are in WT. As per our hypothesis, there should be a strong interaction between the galacturonic acid containing pectin and lignin of the wildtype switchgrass due to which the aggregates cannot be formed as fast as the galacturonic acid lacking GAUT4-kd switchgrass. Biswal et al report that the hot water pretreated GAUT4-kd switchgrass release more sugar after enzymatic hydrolysis than wildtype due to greater cellulose accessible surface.[24] Based on our SANS results, as more lignin is re-distributed in GAUT4-kd it would have led to increased cellulose accessibility due to which an increase in sugar release post pretreatment in GAUT4-kd switchgrass was measured.

Pectin content in general accounts for less than 10% of the total polymers present in the whole plant cell wall.[30] Out of all the pectin polymers, homogalacturonan is present in the largest amount in the cell wall (~80%).[17] Galacturonic acid forms the backbone of homogalacturonan and it is also present in other pectin polysaccharides such as rhamnogalacturonan I but in minor quantities. Our

SANS study indicates a strong interaction between lignin and galacturonic acid containing pectin. However, homogalacturonan is a linear polymer which would give less opportunity for incoming lignin monomers to covalently bond with the galacturonic acid.

To test if a linear polymer like homogalacturonan can make a bond with lignin, we made model composite by polymerizing monolignol in presence of a solution containing homogalacturonan. The x-ray scattering of the DHP-HG composite is different from both the HG-only and DHP-only scattering profile and is indicative of DHP forming aggregates in a HG network. The presence of DHP with the HG even after centrifugation shows that at least a physical interaction exists between DHP and HG. Based on the scattering profile, a schematic showing DHP in HG solution is shown in Figure 6.12. Further, the IR spectra of the DHP-HG composite shows a unique band that is not present in the DHP only and HG only sample. Such a band is also not present in the IR spectra obtained from mixing the DHP and HG together. This band occurs in the region in which ester bonds are often seen. The DHP is formed by a free radical reaction in which the enzyme horseradish peroxidase in presence of hydrogen peroxide catalyzes the monolignol to free radical or a quinone methide intermediate. It has been proposed that nucleophilic groups such as carboxylic acid may re-aromatize the quinone methide intermediate to form ether or ester linkages respectively at the β position of the monolignol. From an IR spectrum it is not conclusive to say the location at which the possible ester bond would have formed. However, due to the presence of a band in the ester region of the HG-DHP composite the IR spectra does provide preliminary evidence for a chemical interaction and future experiments with nuclear magnetic resonance should strengthen this finding.

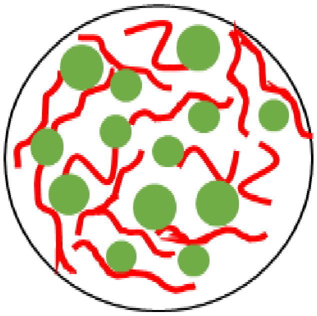


Figure 6.12: Schematic showing the lignin in aggregated state entrapped in homogalacturonan containing solution

6.5). Conclusion and future work

SANS of the GAUT4-kd and wildtype switchgrass both before and after pretreatment enabled us to look at the changes that occur in the cell wall of these variants as a result of hot water pretreatment. No difference in the overall arrangement of the cell wall polymers in GAUT4-kd and wildtype switchgrass were visible with SANS. However, post pretreatment the observation that the lignin aggregates being both greater in size and more in number in GAUT4-kd switchgrass than wildtype is suggestive of a strong interaction between galacturonic acid-containing pectin and lignin. We further investigated the interaction by synthetizing lignin-pectin composites. The lignin-like polymer (DHP) was synthesized in presence of homogalacturonan containing solution. The x-ray scattering curve of such composite was different from a solution in which the DP and homogalacturonan were physically mixed. The composite showed DHP to be trapped in and forming aggregates within a HG network. Our scattering results of the composites also suggests a physical interaction between lignin and homogalacturonan. We wanted to test if the interaction of acid group containing homogalacturonan and lignin is also chemical in nature - as previously it has been shown that small molecules such as acetic acid and synthetic lignin can form an ester linkage.[9] We used infrared spectroscopy and observed a unique ester band only existing in the lignin- homogalacturonan composite. However, since IR is not a conclusive way to determine the linkage and is just gives preliminary evidence, we need to further confirmation if the covalent bond exists with nuclear magnetic resonance spectroscopy.

6.5.1). Nuclear magnetic resonance of lignin-pectin composites

Nuclear magnetic resonance (NMR) spectroscopy of lignin-pectin composite can be done to demonstrate if lignin and pectin (homogalacturonan) are linked by a covalent bond. In a recent study, lignin was synthesized using coniferyl alcohol and horseradish peroxidase as the oxidant in the presence of galacturonic acid containing solution. Using heteronuclear single quantum coherence (HSQC), they were able to observe a unique signal (at 74.8/5.92 ppm) which was assigned to benzyl ester linkage between the carboxylic acid of galacturonic acid and alpha carbon of coniferyl alcohol (Figure 6.13).[9]

It was also reported that as galacturonic acid is a small molecule, the signal for the particular linkage was low and it was difficult to confirm the linkage with other 2D NMR techniques such as HMBC (Heteronuclear Multiple Bond Correlation) or HSQC-TOCSY (Total Correlated Spectroscopy). Homogalacturonan is a galacturonic acid containing polymer and it would give a better signal than galacturonic acid. It is probable that a benzyl ester linkage might exist in the lignin-homogalacturonan composite.

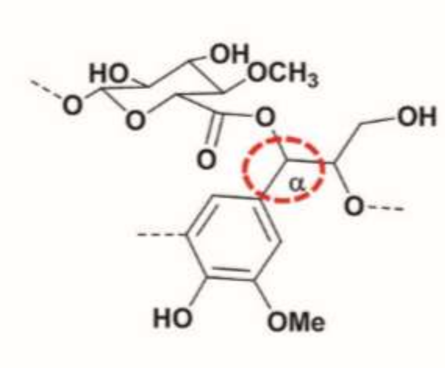


Figure 6.13: Benzyl ester linkage between the carboxylic acid containing galacturonic acid and the Cα carbon of coniferyl alcohol[9]

6.7). References

1. Zhao, X.; Zhang, L.; Liu, D., Biomass recalcitrance. Part I: the chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. *Biofuels, Bioproducts and Biorefining* **2012,** *6* (4), 465-482.
2. Jung, Y. H.; H., K. K., Chapter 3. Acidic Pretreatment. In *Pretreatment of Biomass*, Pandey, A.; Negi, S.; Binod, P.; Larroche, C., Eds. Elsevier: Amsterdam, 2015; pp 259-264.
3. Yang, B.; Dai, Z.; Ding, S.-Y.; Wyman, C. E., Enzymatic hydrolysis of cellulosic biomass. *Biofuels* **2011,** *2* (4), 421-449.
4. Yang, J.; Zhang, X.; Yong, Q.; Yu, S., Three-stage enzymatic hydrolysis of steam-exploded corn stover at high substrate concentration. *Bioresource Technology* **2011,** *102* (7), 4905-4908.
5. Sannigrahi, P.; Ragauskas, A. J.; Miller, S. J., Effects of Two-Stage Dilute Acid Pretreatment on the Structure and Composition of Lignin and Cellulose in Loblolly Pine. *BioEnergy Research* **2008,** *1* (3), 205-214.
6. Donohoe, B. S.; Decker, S. R.; Tucker, M. P.; Himmel, M. E.; Vinzant, T. B., Visualizing lignin coalescence and migration through maize cell walls following thermochemical pretreatment. *Biotechnology and Bioengineering* **2008,** *101* (5), 913-925.
7. DeMartini, J. D.; Pattathil, S.; Avci, U.; Szekalski, K.; Mazumder, K.; Hahn, M. G.; Wyman, C. E., Application of monoclonal antibodies to investigate plant cell wall deconstruction for biofuels production. *Energy & Environmental Science* **2011,** *4* (10), 4332-4339.
8. Erdmann, J., Ueber die Concretionen in den Birnen. *Justus Liebigs Annalen der Chemie* **1866,** *138* (1), 1-19.
9. Giummarella, N.; Pu, Y.; Ragauskas, A. J.; Lawoko, M., A critical review on the analysis of lignin carbohydrate bonds. *Green Chemistry* **2019,** *21* (7), 1573-1595.
10. Tarasov, D.; Leitch, M.; Fatehi, P., Lignin-carbohydrate complexes: properties, applications, analyses, and methods of extraction: a review. *Biotechnology for biofuels* **2018,** *11*, 269-269.
11. Aminzadeh, S.; Zhang, L. M.; Henriksson, G., A possible explanation for the structural inhomogeneity of lignin in LCC networks. *Wood Science and Technology* **2017,** *51* (6), 1365-1376.
12. Balakshin, M.; Capanema, E.; Gracz, H.; Chang, H. M.; Jameel, H., Quantification of lignin-carbohydrate linkages with high-resolution NMR spectroscopy. *Planta* **2011,** *233* (6), 1097-1110.
13. Lawoko, M.; Henriksson, G.; Gellerstedt, G., Structural differences between the lignin-carbohydrate complexes present in wood and in chemical pulps. *Biomacromolecules* **2005,** *6* (6), 3467-3473.
14. Lawoko, M., Unveiling the structure and ultrastructure of lignin carbohydrate complexes in softwoods. *Int J Biol Macromol* **2013,** *62*, 705-713.
15. Giummarella, N.; Zhang, L.; Henriksson, G.; Lawoko, M., Structural features of mildly fractionated lignin carbohydrate complexes (LCC) from spruce. *RSC Advances* **2016,** *6* (48), 42120-42131.
16. Cosgrove, D. C., Comparative structure and biomechanics of plant primary and secondary cell walls. *Frontiers in plant science* **2012,** *3*, 204.
17. Albersheim, P.; Darvill, A.; Roberts, K.; Sederoff, R.; Staehelin, A., *Plant Cell Walls: From Chemistry to Biology*. Garland Science: 2011.

18. Grabber, J. H.; Ralph, J.; Hatfield, R. D.; Quideau, S.; Kuster, T.; Pell, A. N., Dehydrogenation Polymer−Cell Wall Complexes as a Model for Lignified Grass Walls. *Journal of Agricultural and Food Chemistry* **1996,** *44* (6), 1453-1459.

19. Grabber, J. H.; Hatfield, R. D.; Ralph, J.; Zoń, J.; Amrhein, N., Ferulate cross-linking in cell walls isolated from maize cell suspensions. *Phytochemistry* **1995,** *40* (4), 1077-1082.

20. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nature Communications* **2019,** *10* (1), 347.

21. Zykwinska, A.; Thibault, J.-F.; Ralet, M.-C., Competitive binding of pectin and xyloglucan with primary cell wall cellulose. *Carbohydrate Polymers* **2008,** *74* (4), 957-961.

22. Voiniciuc, C.; Pauly, M.; Usadel, B., Monitoring Polysaccharide Dynamics in the Plant Cell Wall. *Plant physiology* **2018,** *176* (4), 2590-2600.

23. Minor, J. L., Location of Lignin-Bonded Pectic Polysaccharides. *Journal of Wood Chemistry and Technology* **1991,** *11* (2), 159-169.

24. Biswal, A. K.; Atmodjo, M. A.-O.; Li, M. A.-O.; Baxter, H. L.; Yoo, C. G.; Pu, Y. A.-O.; Lee, Y. C.; Mazarei, M.; Black, I. M.; Zhang, J. Y.; Ramanna, H.; Bray, A. L.; King, Z. R.; LaFayette, P. R.; Pattathil, S.; Donohoe, B. A.-O.; Mohanty, S. S.; Ryno, D.; Yee, K.; Thompson, O. A.; Rodriguez, M., Jr.; Dumitrache, A.; Natzke, J.; Winkeler, K.; Collins, C.; Yang, X. A.-O.; Tan, L.; Sykes, R. W.; Gjersing, E. L.; Ziebell, A.; Turner, G. B.; Decker, S. A.-O.; Hahn, M. A.-O.; Davison, B. H.; Udvardi, M. K.; Mielenz, J. R.; Davis, M. A.-O.; Nelson, R. S.; Parrott, W. A.; Ragauskas, A. J.; Neal Stewart, C., Jr.; Mohnen, D., Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. **2018,** *36* (3), 249-257.

25. Li, Y.; Fu, Q.; Yu, S.; Yan, M.; Berglund, L., Optically Transparent Wood from a Nanoporous Cellulosic Template: Combining Functional and Structural Performance. *Biomacromolecules* **2016,** *17* (4), 1358-1364.

26. Muraille, L.; Pernes, M.; Habrant, A.; Serimaa, R.; Molinari, M.; Aguié-Béghin, V.; Chabbert, B., Impact of lignin on water sorption properties of bioinspired self-assemblies of lignocellulosic polymers. *European Polymer Journal* **2015,** *64*, 21-35.

27. Haitao Yang; Xing Zheng; Lan Yao; Ruijin Gu; Xie, Y., Elucidation of the bonds between cellulose and dehydrogenation polymer with carbon-13 isotopic tracer method. *Cellulose Chemistry and Technology* **2013,** *47* (3-4), 143-151.

28. Hammouda, B.; Ho, D. L., Insight into chain dimensions in PEO/water solutions. *Journal of Polymer Science Part B: Polymer Physics* **2007,** *45* (16), 2196-2200.

29. Pingali, S. V.; Urban, V. S.; Heller, W. T.; McGaughey, J.; O'Neill, H.; Foston, M. B.; Li, H.; Wyman, C. E.; Myles, D. A.; Langan, P.; Ragauskas, A.; Davison, B.; Evans, B. R., Understanding Multiscale Structural Changes During Dilute Acid Pretreatment of Switchgrass and Poplar. *ACS Sustainable Chemistry & Engineering* **2017,** *5* (1), 426-435.

30. Voragen, A. G. J.; Coenen, G.-J.; Verhoef, R. P.; Schols, H. A., Pectin, a versatile polysaccharide present in plant cell walls. *Structural Chemistry* **2009,** *20* (2), 263.

31. Garvey, C. J.; Knott, R. B.; Searson, M.; Conroy, J. P., USANS study of wood structure. *Physica B: Condensed Matter* **2006,** *385-386*, 877-879.

Chapter 7). Conclusion and future outlook

7.1). Conclusion

The structure of the plant cell wall is due to the arrangement of the polymers within the cell wall and the interactions between these polymers. In my thesis, aspects of the impact that polymer interactions have in defining cell wall architecture were investigated. New insights (as summarized in Figure 7.1) into these interactions were obtained by either using model cell wall mimicking composite systems or studying their structural features in intact biomass and during deconstruction. The use of small-angle neutron scattering (SANS) with contrast variation and specialized sample environments enabled us to study changes in the structure of individual polymers and in the overall cell wall due to hot water and dilute acid pretreatment. The results of this work are of value for engineering cell walls with reduce recalcitrance and improving pretreatment approaches.

One of the major findings of this work was to learn that changes in cellulose structure post dilute acid pretreatment were dependent on the type of hemicellulose backbone and structure. For this work, model hemicellulose-cellulose composites were prepared by synthesizing bacterial cellulose in presence of glucomannan or xyloglucan dissolved in the growth media. SANS was used to study structural changes in the composites as a result of dilute acid pretreatment (DAP). By growing deuterium labeled cellulose in the presence of hydrogenated hemicellulose it was possible to deconvolute the scattering signatures of the two components. No significant change in the crystallinity and glucan chain packing was observed in the DAP cellulose by X-ray diffraction (XRD).

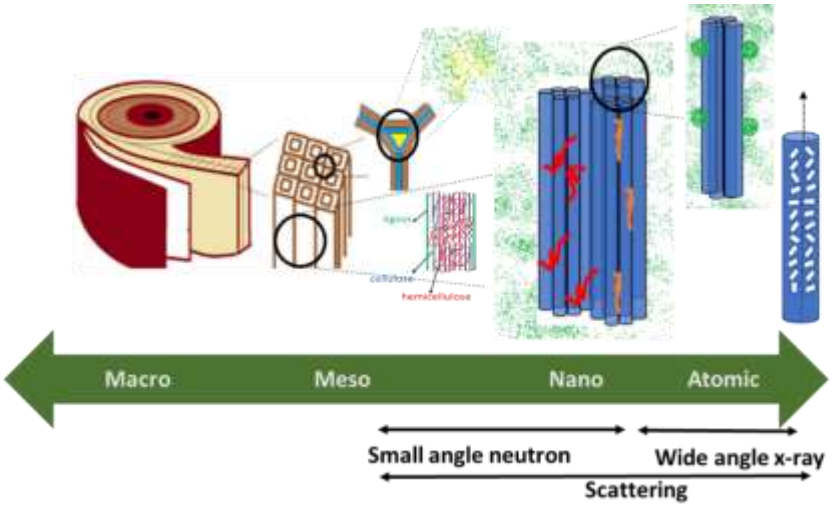


Figure 7.1: Small angle neutron scattering and x-ray diffraction were used to probe the different length spanning several nanometers to a few angstroms and different insights into cell wall polymer interactions were gained. Model hemicellulose-cellulose composites were synthesized by growing bacterial cellulose in media containing different hemicelluloses (xyloglucan shown in orange and glucomannan shown in red). Using scattering, it was found the different hemicellulose changed cellulose structure differently post dilute acid pretreatment. This result shows that interaction of hemicellulose with cellulose (shown in blue) is dependent on the backbone and side chain structure of hemicellulose. Two naturally occurring poplar variant were studied with scattering to probe for structural differences as the low lignin containing poplar realized more glucan than high lignin containing poplar both before and after hot water pretreatment. Lignin was freely distributed with other cell wall polymers in high lignin containing variant while it formed aggregates in low lignin containing variants. Post pretreatment the crystallite within the cellulose microfibril (blue) were better aligned in low lignin containing poplar than high lignin containing poplar. This result suggests that as polymers are getting deposited into the cell wall, they can influence changes in each other structures which can impact their breakdown by enzymes. Finally, we explored the interactions between lignin and homogalacturonan (yellow). Our results suggest that an interaction exists between these two polymers. Proposed future experiments will allow us able to conclude if these interations are physical or chemical in nature.

However, at the nanoscale, DAP cellulose showed a collapse in structure as indicated by a decrease in R_g from 250Å to 130 Å and a change of power law exponent (alpha) from 2.8 to 3.1. This was interpreted as expulsion of water from the space between microfibrils resulting in formation of a tightly packed macrofibril.

In addition, the cellulose network after DAP show less entanglement than in the untreated cellulose. Addition of xyloglucan (XG) decreased cellulose crystallinity in the composite material and altered the glucan chain packing in the microfibrils as determined by analysis of the XRD pattern. After removal of the XG by DAP, the nanoscale cellulose structure showed minimal change. On the other hand, in composites formed using glucomannan (GM) the structural changes that occurred due to DAP were similar to the cellulose only samples. However, the cellulose network of the pretreated composite was more entangled (α= 2.8) than the cellulose only counterpart (α=2.2). Our results show that XG and GM interact with the cellulose network differently. XG directly interacts with the microfibrils that comprise the macrofibrils, while GM only interacts at the surface of the macrofibrils.

For the next study we took a comparative approach to learn about interaction between lignin and cellulose. Structural changes in the cell wall of two naturally occurring variants of poplar species *Populus trichocarpa*, BESC-316 and GW-11012 with Klason lignin contents of 17.8 and 23.2% respectively, were investigated both before and after hot water pretreatment. It was previously shown that the lower lignin content genotype showed greater sugar release before and after hot water pretreatment compared to the lower lignin counterpart. We used small-angle neutron scattering (SANS) and wide-angle x-ray scattering (WAXS) to investigate the structural changes in the BESC-316 and GW-11012 genotypes that were subjected to hot water pretreatment (160˚C 70min, 180˚C 18min and 180˚C 45min) to obtain information about cellulose microfibril organization, the lignin and hemicellulose network, and overall morphology of the plant cell walls. Cellulose microfibril arrangement in GW-11012 is consistent with aggregated microfibrils and differed significantly from the well-ordered cellulose microfibrils in BESC-316 before pretreatment. Post-pretreatment, little change was seen in cellulose arrangement for GW-11012 whereas BESC-316 showed aggregation of microfibrils. SANS showed that GW-11012 had increased scattering intensity in the mid Q region compared to BESC-316. After pretreatment, both genotypes have very similar scattering patterns indicative of similar structural changes occurring in the pretreated cell walls. Cellulose accessibility was measured using the modified Simons' stain and showed that GW-11012 had greater accessibility compared to BESC-316. However, post pretreatment similar values were obtained for GW-11012 and BESC-316. No significant differences in the amount of 5-hydroxymethyl furfural, furfural and acetic acid for BESC-316 and GW-11012 were observed in the pretreatment liquor. The crystallite size and orientation where both larger and better arranged in GW-11012 than BESC-316 and this trend was maintained post pretreatment. This study showed that deposition of polymers in the cell wall is co-dependent, as structural changes both in lignin distribution and in cellulose crystallite size and orientation were simultaneously observed. Such changes may contribute in cellulose surface accessible for enzymatic hydrolysis.

Finally, lignin and pectin interactions were probed by comparing the lignin aggregates in hot water pretreated wild type and galacturonosyl transferase (GAUT4) knock-down switchgrass and by making lignin-homogalacturonan (HG) composites. The GAUT4 gene is responsible for making the enzyme that links two galacturonic acid residues by α (1-4) linkages. The GAUT4-kd showed reduced recalcitrance than the wildtype and this was associated to the less crosslinked pectic network of the GAUT4-kd as compared to the wildtype. This was surprising since pectin is a minor cell wall component accounting for less than ten percent of the total cell wall composition. We investigated if reasons other the less entangled pectic network are responsible for reduced recalcitrance of GAUT4-kd switchgrass. We studied the structural differences in the cell wall of GAUT4-kd and wildtype switchgrass using SANS. There was no difference in the cellulose microfibril cross section or the degree of entanglement of the polymers in the cell wall matrix. However, hot water pretreated GAUT4-kd showed a slight difference in the lignin aggregate size and number than the wildtype. The difference in aggregate size indicates an association between lignin and galacturonan containing polysaccharides. To further investigate the nature of this association we made lignin-HG composites as majority of the pectin present in the cell wall is HG and studied the composites with SAXS and IR. The composites were made by polymerizing lignin like dehydrogenation polymer in presence of HG. The SAXS curve of the composites showed a network like structure in which the DHP aggregates were present along with the HG. The IR spectra of the composites showed a unique band in the ester region that was present only in the composites and not in the homogalacturonan, DHP or physical mixture of homogalacturonan and DHP. A combination of the SANS results on the post hot water pretreated GAUT4-kd and wildtype switchgrass and SAXS and IR study of the DHP-homogalacturonan composite suggests a covalent linkage between lignin and homogalacturonan. In the absence of homogalacturonan such a linkage would not form and along with less crosslinked pectin network, reduced recalcitrance was observed for GAUT4-kd switchgrass than the wildtype.

In summary, through these three studies we examined individual cell wall polymers and cell wall of feedstocks such as poplar and switchgrass with dilute acid and hot water pretreatment to learn about the different polymer interactions and the changes in these post pretreatments. These include changes cellulose structure post pretreatment are dependent on the backbone and side chain composition of hemicellulose, besides in plant cell wall polymer compositions their arrangement in the cell wall can also influence enzymatic hydrolysis and studied the interactions between homogalacturonan and lignin.

7.2). Future outlook

Research spanning many decades have been invested in understanding the structure of plant cell walls. Today we are able to understand the structure much better, however, a predictive understanding is not yet complete. From just examining the ultrastructure of cellulose, there are several remaining fundamental questions such as the number of chains and shape of the cellulose microfibrils, factors that make part certain region of cellulose crystallite and others amorphous, the arrangement of amorphous and crystalline part in the cellulose microfibril and if crystallinity affects enzyme hydrolysis which are not known with absolute certainty. Besides making cellulosic based ethanol another use of cellulose is in making cellulose nanocrystals, and knowledge about the structure of cellulose may be helpful in chemically modifying it and making nanocellulose with desired properties. One such example could be tailoring the properties of nanocellulose such that it can transported in dried form and upon rehydration it still has the same structure and properties.

The hemicelluloses are at times acetylated and degree of acetylation is variable. They also have differing degree of branching and type of sugars in the branches. Hemicellulose and cellulose are two major cell wall components that interact with each other. Using the same approach of growing bacterial cellulose in presence of different hemicellulose, as in Chapter 5 of this thesis, it would be worth investigating how the degree of acetylation and branching affect the ultrastructure of cellulose. Depending in whether the acetyl groups or branches change the cellulose ultrastructure the enzyme hydrolysis of such cellulose may positively or negatively hydrolysis. The information may be useful in engineering plants with reduced recalcitrance.

Studying the two natural variants with SANS showed a difference in organization of cellulose and lignin and it is probable that deposition of lignin impacted the changes in the cellulose ultrastructure. These two variants are a part of the GWAS done on about 1100 poplars and to build on the study of Chapter 5, the promising low recalcitrant poplar should be studied to see if structural differences are responsible for the observed low recalcitrance. Additionally, lignin has become increasing important from the standpoint of making value added products and hence studying lignin structure becomes important for processing it to eventually convert into such products. The naturally occurring variants of poplar have varying ratios of lignin and different monolignol composition. They are a valuable resource to study how the differences in S:G ratio and linkages affects lignin overall conformation in different solvents.

Overall, research in the fields of cell wall formation and deconstruction of cell wall whether for biofuel or bioproducts go hand and hand. Using the knowledge of both the fields, a deeper understanding of the cell wall models plants can be obtained. Different chemical and molecular features of the cell wall should be investigated with already known and with constantly developing novel characterizations methods. Also, using a multi length scale approach complementary information can be obtained that if pieced together can help solve the puzzle - the structure of plant cell wall.

Appendix

A). Interaction between cellulose III and enzyme, cellobiohydrolyase

A.1). Introduction

Several pretreatment methods like hydrothermal, dilute acid, lime, ammonia percolation and organosolv have been used to detach the cellulose and hemicellulose from other biomass components such as lignin to improve subsequent enzymatic hydrolysis.[1] However, our understanding of how enzymes interact with the pretreated biomass from these various methods is limited. The major biomass components cellulose, hemicellulose and lignin share similar scattering length densities, which are also close to those of the hydrolytic enzymes. This makes it difficult to process the neutron scattering data for the whole biomass. Hence, we use bacterial cellulose as a model system because it is free from hemicellulose and lignin and can be deuterated. Contrast matching of reaction mixtures of deuterated cellulose and protiated enzymes could enable us to see the interaction of enzyme with cellulose and structural changes of cellulose both during and after enzyme hydrolysis. Our group has reported interactions of cellulose with Cel7A using neutron scattering and molecular simulations.[2] Now we intend to pretreat the deuterated cellulose with the ammonia method and check the interactions of Cel7A with this pretreated cellulose. It has been reported that pretreatment of cellulose with ammonia, which changes the crystalline form of the cellulose from the native cellulose I to cellulose III,[3] gives a better yield of total sugars and avoids formation of inhibitory products that might interfere with hydrolytic enzymes.[4]

Our aim was to understand how the industrially relevant fungal enzyme cellobiohydrolase (Cel7A) hydrolyzes cellulose III which is formed by treating cellulose with ammonia. Small angle neutron scattering would enable us to leverage the difference in scattering power between protiated Cel7A and deuterated cellulose III and enable us to visualize the conformations of Cel7A during Cellulose III digestion.

A.2). Results

We have successfully purified the enzyme Cel7A from commercially obtained culture filtrate from the fungus *Trichoderma reesei* using a published method.[2] The enzyme was purified using gel filtration and chromatofocusing chromatography. Enzyme preparations were characterized by SDS-PAGE and isoelectric focusing. A single band at 67 kDa was observed on SDS-PAGE and a band at pI of 4.2 was seen on IEF gel (Figure A1), corresponding to the reported values for this enzyme. Circular dichroism (CD) spectroscopy confirmed that the purified protein had a well folded protein structure. We have successfully prepared deuterated cellulose III from deuterated bacterial cellulose using a reported method,[3] confirming the crystalline form by X-ray diffraction (Figure A2). We also studied the hydrolysis of native bacterial cellulose and cellulose III with enzyme Cel7A. As expected, we saw that Cel7A hydrolyzed cellulose III faster and produced more of the product cellobiose than native bacterial cellulose (cellulose I).[4]

SANS curve of deuterated cellulose III in 85% D_2O still had some residual scattering from cellulose (Figure A3). The scattering pattern collected for Cel7A bound to cellulose III looks identical in sodium acetate buffer pH 4.2 and 5 but a slight difference in the fall of the scattering pattern is seen at pH.7 (Figure A4). The data analysis is in progress however, the residual scattering is overpowering the highQ and making it difficult to conclusively determine the differences in the conformation of bound Cel7A.

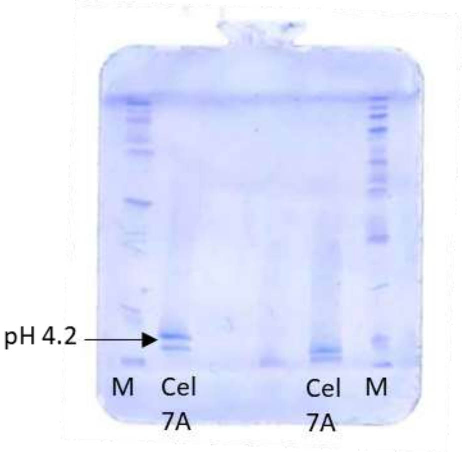


Figure A3: Isoelectric focusing gel showing separation of the cellobiohydrolases with different glycosylation. The sample was passed through the isoelectric focusing resin to separate the cellobiohydrolases isoforms and the top band was used for the SANS study

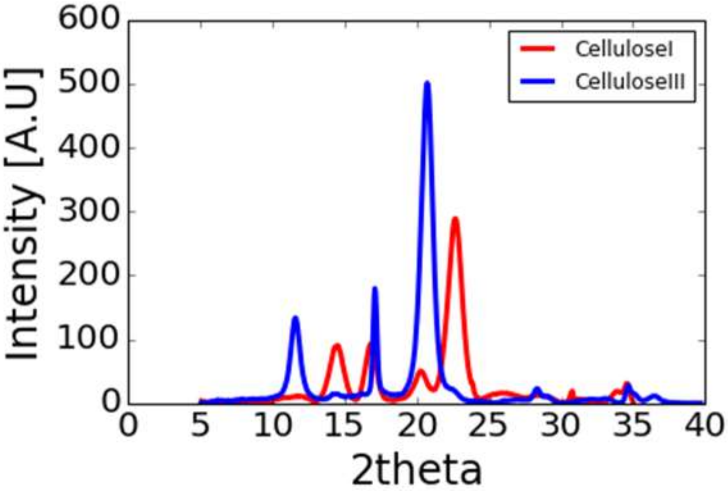


Figure A4: We prepared deuterated cellulose and used a pressure vessel to first liquify the ammonia and soak the sample and then to heat the sample to 140˚C at which the cellulose I is converted to cellulose III. The formation of cellulose III was confirmed by taking a diffraction pattern and a comparison to the cellulose I in shown

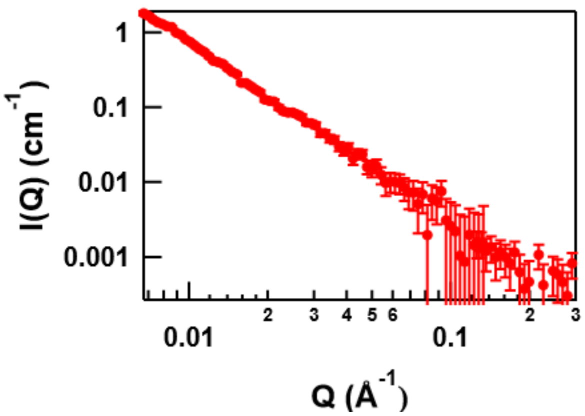


Figure A5: SANS curve for partially deuterated cellulose III in 85% D_2O

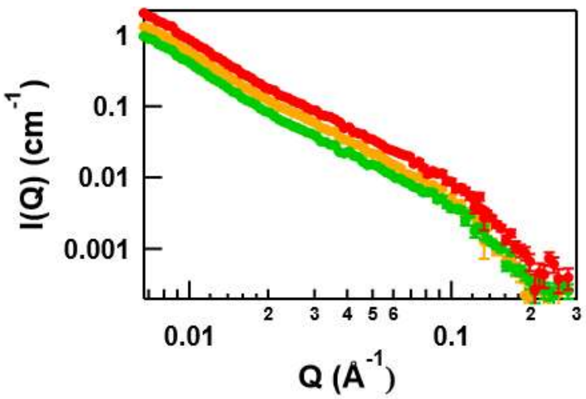


Figure A6: SANS curve for Cel7A bound to cellulose III in sodium acetate buffer with pH4.2 (red), pH 5.0 (orange) and pH 7.0 (green)

B). Method development for bacterial cellulose

B.1). Deuterium level incorporation using infrared spectroscopy

The cellulose labelled with deuterium (D) had a different scattering length density than hydrogenated hemicellulose. This leveraged us to see specific changes in the structure of cellulose in presence of different hemicellulose with neutron scattering. In order to determine the contrast match point, we synthesized bacterial cellulose in different level of deuterium containing media. To quantitate the level of D incorporation we measured the infrared spectra of each of the freeze-dried bacterial cellulose pellicle. As the D-level in the media kept increasing, a trend in the IR spectra with decreasing -CH peak and increasing -CD peak was observed (Figure B1 panel A). All the IR spectra were corrected for the background and the -CH peak was fitted with a gaussian function using the peak fitting macros in the software Igor. The peak height of the -CH peak was plotted against the D-level of the media (Figure B1 panel B). The purpose of this method development was quality assurance and making sure that D-level in the pellicle is always such that the partially deuterated cellulose can be matched with 85% D2O in neutron scattering experiment. This would avoid the need to perform contrast variation of bacterial cellulose each time it is used.

B.2). Oriented bacterial cellulose

The scanning electron microscopy images of freeze dried bacterial cellulose show that the ribbon like structure of cellulose fibers are randomly arranged.[5] We wanted to test if an external force along the axial direction if applied could lead to aligning the cellulose ribbons and the crystallites within the ribbons. Figure B2 shows the setup used for aligning bacterial cellulose and Figure B3 shows x-ray diffraction images of aligned bacterial cellulose obtained using this setup.

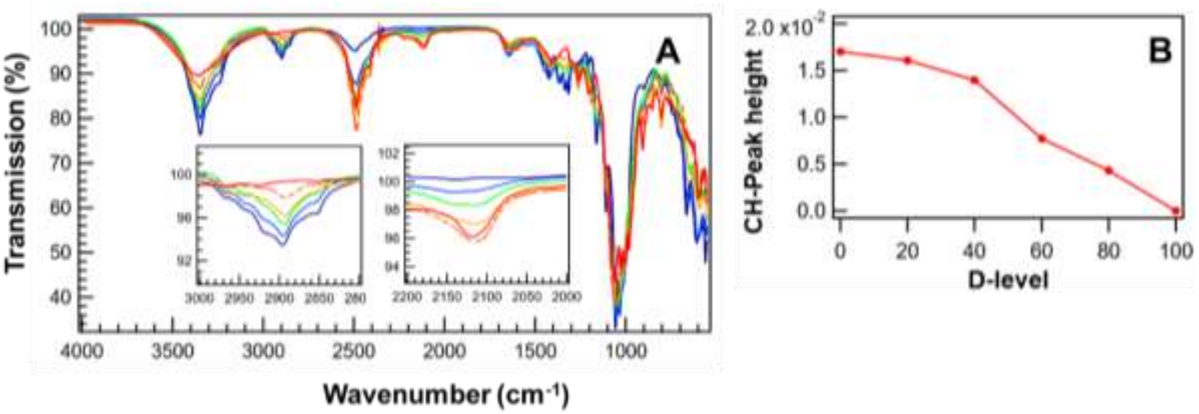


Figure B3: panel A shows the infrared spectra of the different D-level containing freeze-dried bacterial cellulose. The -CH region (2800 to 3000 cm^{-1}) and -CD region (2000 to 2200 cm^{-1}) are shown enlarged in the inset. Panel B shows a plot of CH peak height determined by fitting the -CH peak of the infrared spectra of differently labelled bacterial cellulose against the D-level in which the bacterial cellulose was grown.

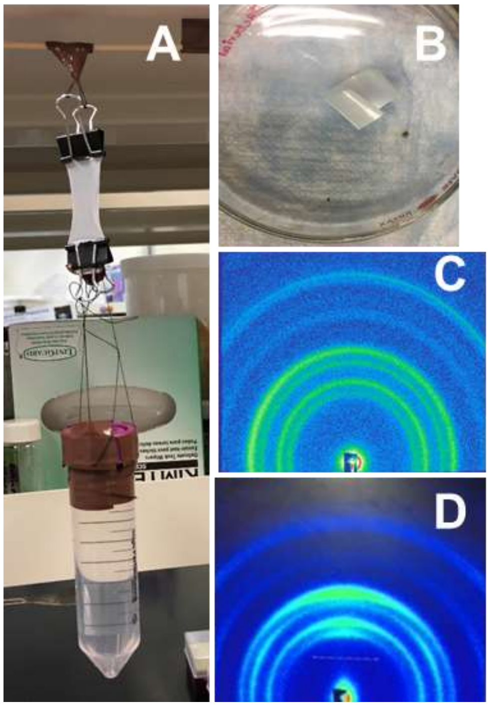


Figure B4: Panel A shows the setup which was used to align bacterial cellulose and panel B shows bacterial cellulose which was dried on a petri-dish and used as a control. Panel C and D are the 2 D x-ray diffraction pattern observed of the aligned and non-aligned bacterial cellulose respectively

References

1. Mosier, N.; Wyman, C.; Dale, B.; Elander, R.; Lee, Y. Y.; Holtzapple, M.; Ladisch, M., Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology* **2005,** *96* (6), 673-686.
2. Pingali, S. V.; O'Neill, H. M.; McGaughey, J.; Urban, V. S.; Rempe, C. S.; Petridis, L.; Smith, J. C.; Evans, B. R.; Heller, W. T., Small Angle Neutron Scattering Reveals pH-dependent Conformational Changes in Trichoderma reesei Cellobiohydrolase I: implications for enzymatic activity. *Journal of Biological Chemistry* **2011,** *286* (37), 32801-32809.
3. Wada, M.; Chanzy, H.; Nishiyama, Y.; Langan, P., Cellulose IIII Crystal Structure and Hydrogen Bonding by Synchrotron X-ray and Neutron Fiber Diffraction. *Macromolecules* **2004,** *37* (23), 8548-8555.
4. Igarashi, K.; Wada, M.; Samejima, M., Activation of crystalline cellulose to cellulose III_I results in efficient hydrolysis by cellobiohydrolase. *The FEBS Journal* **2007,** *274* (7), 1785-1792.
5. Martinez-Sanz, M.; Gidley, M. J.; Gilbert, E. P., Application of X-ray and neutron small angle scattering techniques to study the hierarchical structure of plant cell walls: a review. *Carbohydr Polym* **2015,** *125*, 120-34.

Vita

The author was born in Ahmedabad, Gujarat, India to Darshana Shalibhadra Shah and Shalibhadra Shantilal Shah. She attended Vyavasayi Vidya Pratishthan (V.V.P) engineering college of Saurashtra university and graduated in 2009 with a bachelor's degree in biotechnology engineering. She then chose to pursue further education in her area of interest of fermentation science and biochemical engineering. She was accepted in a master's degree program in chemical and biochemical engineering at University of California, Irvine (UCI). She graduated from UCI with a Master of Science in 2012. After this, she got her first job and worked at the Oak Ridge National Laboratory (ORNL) as a post masters' research associate for two years. During this time, she realized her passion for research. She was accepted for a doctoral degree program at the Bredesen Center of Interdisciplinary Science and Research at University of Tennessee, Knoxville (UTK). She completed her coursework at UTK and did research for her doctoral dissertation at ORNL. She joined the program in January of 2015, has defended her dissertation and is expected to complete the requirements of the degree program in January 2020. The author's research interest is in understanding structure of biological materials to improve their function or properties and her career goal is to continue research in this area.