University of Kentucky

UKnowledge

2023

APPLICATION OF MASS SPECTROMETRY FOR THE CHARACTERIZATION OF SYNTHETIC OLIGOMERS AND NATURAL LIGNIN

Poorya Kamali
University of Kentucky, poorya.kamali@gmail.com
Author ORCID Identifier:
https://orcid.org/0009-0009-3315-1861
Digital Object Identifier: https://doi.org/10.13023/etd.2023.198

Right click to open a feedback form in a new tab to let us know how this document benefits you.

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Poorya Kamali, Student

Dr. Bert C. Lynn, Major Professor

Dr. Dong-Sheng Yang, Director of Graduate Studies

APPLICATION OF MASS SPECTROMETRY FOR THE CHARACTERIZATION OF
SYNTHETIC OLIGOMERS AND NATURAL LIGNIN

_____

DISSERTATION
_____

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Art and Sciences
at the University of Kentucky

By
Poorya Kamali
Lexington, Kentucky
Director: Dr. Bert C. Lynn, Professor of Chemistry
Lexington, Kentucky
2023

ABSTRACT OF DISSERTATION

APPLICATION OF MASS SPECTROMETRY FOR THE CHARACTERIZATION OF
SYNTHETIC OLIGOMERS AND NATURAL LIGNIN

As part of the ongoing effort to substitute finite fuel and chemical resources with renewable ones, biomass is emerging as one of the most promising sources. Biomass consists of three main components of cellulose, hemicellulose, and lignin. Traditionally, cellulose has been used extensively in pulping industry, while lignin has been considered waste and is burned to generate heat. Lignin, a complex aromatic polymer component of biomass, has the potential to be used as a source of aromatic chemicals and pharmaceutical synthons. The recalcitrant nature of lignin, the lack of effective lignin breakdown methods and analytical techniques to analyze it are the main obstacles to obtaining high-yield chemicals from lignin. Mass spectrometry has proven to be one of the most promising analytical techniques and it is widely used in the pharmaceutical and chemical industries. The goal of this work is to develop analytical methods using mass spectrometry and lignin model compounds. Additionally, this work focused on the development and application of quantitative Derivatization Followed by Reductive Cleavage(q-DFRC) for the evaluation of various biomass pretreatment methods.

Since most commercially available lignin model compounds fail to mimic the structure of native lignin, it is necessary to develop compounds that more closely reflect the functionality of native lignin. The first project of this dissertation is focused on developing precursors for synthesizing β-O-4 model compounds and modifying their functional groups. The precursors have been synthesized and analyzed using gas chromatography-mass spectrometry. These precursors were used to synthesize β-O-4 model compounds that exhibit all characteristics of the native lignin.

The second project involved the synthesis and mass spectral analysis of a mixed linkage trimer containing both β-O-4 and β-5 bond types. A detailed analysis of the mass spectral fragmentation of lignin trimer with lithium adduct ionization is presented. The developed analysis of the lignin trimer facilitates the structural elucidation of lignin breakdown products.

The third project involved the application of q-DFRC as one of the lignin breakdown techniques to evaluate different biomass pretreatment methods. Ethanosolv, dioxosolv, co-solvent enhanced lignocellulosic fractionation (CELF), hydrotropic, and

acetic acid/formic acid pretreatments were evaluated by q-DFRC with deuterium-labeled acylated monolignols internal standard. An evaluation and comparison of the quality of lignin obtained from each of these pretreatments was conducted.

This dissertation provides valuable information for the advancement of mass spectrometric analysis of lignin, and it can be applied to lignin oligomer analysis. Furthermore, the q-DFRC results provide insight into how various pretreatments are related to the extent of condensation in extracted lignin.

KEYWORDS: Lignin, Mass spectrometry, Fragmentation, DFRC, Lignicellulosic biomass pretreatment, Structural analysis

<div align="center">

Poorya Kamali
_____
(Name of Student)

03/24/2023
_____
Date

</div>

This dissertation is dedicated to
Fatemeh and Masseha
My constant source of love and support during this journey

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BSTFA   N,O-Bis(trimethylsilyl)trifluoroacetamide

Calc.   Calculated

CELF   Cosolvent Enhanced Lignocellulosic Fractionation

CIMV   Compagnie Industrielle de la Matière Végétale

CNMR   Carbon Nuclear Magnetic Resonance Spectroscopy

DAD   Diode Array Detector

DC   Direct Current

DIBAL-H   Diisobutylaluminum hydride

DIPEA   Diisopropylethylamine

DMAP   4-Dimethylaminopyridine

EI   Electron Ionization

ESI   ElectroSpray Ionization

FT-ICR   Fourier Transform Ion Cyclotron Resonance mass spectrometer

FT-IR   Fourier Transform InfraRed

GC-MS   Gas Chromatography-Mass Spectrometry

GPC   Gel Permission Chromatography

HCD   High energy Colission Dissociation

HESI   Heated Electrospray Ionization

HMBC   Heteronuclear Multiple Bond Correlation Spectroscopy

HNMR   Hydrogen Nuclear Magnetic Resonance Spectroscopy

HPLC   High Performance Liquid Chromatography

HRAM   High Resolution Accurate Mass

HRMS   High-Resolution Mass Spectrometry

HSQC   Heteronuclear Single Quantum Coherence Spectroscopy

LTQ   Linear Trap Quadrupole Mass Spectrometer

m/z   mass-to-charge ratio

MALDI   Matrix-Assisted Laser Desorption

NCE   Normalized Colission Energy

NMR	Nuclear Magnetic Resonance
Obs.	Observed
p-TsOH	p-Toluenesulfonic acid
ppm	part per million
qDFRC	quantitative Derivatization Followed by Reductive Cleavage
RF	Radio Frequancy
Rf	Retention factor
SEC	Size-Exclusion Chromatography
SPE	Solid Phase Extraction
TEA	Triethylamine
THF	Tetrahydrofuran
TIPS	Triisopropylsilyl group
TLC	Thin Layer Chromatography
TMS	Trimethylsilyl group
XRD	X-Ray Diffraction

CHAPTER 1. INTRODUCTION

1.1   Lignocellulosic Biomass

In recent years, there has been a growing interest in utilizing lignocellulosic biomass as a sustainable and renewable source of energy and chemical feedstocks[1]. Currently, aromatic chemicals are predominantly derived from fossil fuels. However, the finite nature of fossil fuels has made finding alternative sources of these chemicals increasingly critical[2].

Lignocellulosic biomass includes materials such as agricultural residues, forestry waste, and energy crops[3, 4]. It is a complex material, found in plant cell walls, that is composed of three primary components: cellulose, hemicellulose, and lignin[5]. Figure 1.3 shows the schematic view of the biomass.

Cellulose, the most abundant polymer in nature, is a linear polysaccharide consisting of glucose units that are linked by β-1,4-glycosidic bonds. These linear polymer chains are bounded together by strong hydrogen bonds. Cellulose has been widely utilized in the pulp and paper industries and traditionally was the main polymer of interest in biomass[6].

Hemicellulose is a heteropolysaccharide composed of various sugar units that differ depending on the source of the biomass. In contrast to cellulose, hemicellulose is a branched polymer that is more easily broken down into its sugar units[7]. Hemicellulose is the component that links cellulose to lignin, the other components of biomass.

Lignin is a complex, recalcitrant, and crosslinked polymer. It is the most abundant aromatic polymer in nature[1, 8]. Lignin's recalcitrant nature protects cellulose and hemicellulose from enzymatic and chemical degradation[6, 9]. The aromatic nature of lignin

makes it a potentially valuable source for aromatic chemical and pharmaceutical synthons[10].

1.1.1 Lignin

There has been an increasing interest in using lignocellulosic biomass to produce biomaterials and biofuels. Finding an ideal renewable source of energy and chemical has never been as critical as it is in recent years due to finite fossil fuels and an increase in oil prices. In the year 2021, biomass provided 5% of the total energy consumption in the US and it was 10% of the world's energy consumption in 2022[11].

Lignin is the most abundant aromatic polymer in nature, consisting of three main monomers: p-coumaryl, coniferyl, and sinapyl alcohol, which differ in the presence of a methoxy group on the ring[12]. These monolignols when bonded in the polymer are named p-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units[13]. The polymerization of these monomers occurs through oxidative coupling, catalyzed by laccases and peroxidases, leading to the formation of radical monolignols[14, 15]. Due to the delocalization of the radical, lignin polymerizes with various bond types, leading to the complex and intricate structure of this biopolymer.

The most abundant linkages in lignin are the β-O-4 ether bonds, which constitute approximately 50-65% of the linkages in hardwood lignin and 43-50% in softwood lignin, and relatively higher in herbaceous lignin at 74-84%. Other important linkages include the β-5 and β-β linkages, which make up around 3-11% and 3-12% of the linkages in hardwood lignin, respectively. Softwood lignin (2-6%) and herbaceous lignin (1-7%) exhibit a relatively lower abundance of β-β linkages[16-19]. Notably, softwood lignin has

been reported to have β-5 linkages in the range of 9-12%, while herbaceous lignin is ranging from 5-11%[2]. In addition, less frequently observed linkages include 4-O-5 and α-O-4 bond types. The exact composition of the lignin linkages can vary depending on the plant species, tissue type, and growth conditions[20].

The nomenclature for the various types of lignin bonds is based on the position of the monomeric unit on the ring and the α, β, and γ carbons on the aliphatic chain (as depicted in Figure 1.1) which are forming the bond. For example, if a monomeric unit is linked from the β position to the phenolic oxygen on position 4 of the ring of the other monomeric unit, the bond type is called β-O-4. The natural structure of lignin is complex and cross-linked with various bond types, although its exact structure is not fully known. A proposed simplified structure of lignin is shown in Figure 1.2, illustrating the different types of linkages.

Various analytical techniques are used to determine the composition, structure, and properties of lignin. For example, nuclear magnetic resonance (NMR) spectroscopy is commonly used to identify and quantify different types of chemical bonds in lignin[21, 22]. Fourier transform infrared (FTIR) spectroscopy is another technique that provides information about the functional groups present in lignin[23]. Gel permeation chromatography (GPC) and size-exclusion chromatography (SEC) are used to determine the molecular weight distribution of lignin[24-27]. Nonetheless, the precise structure of lignin remains uncertain.
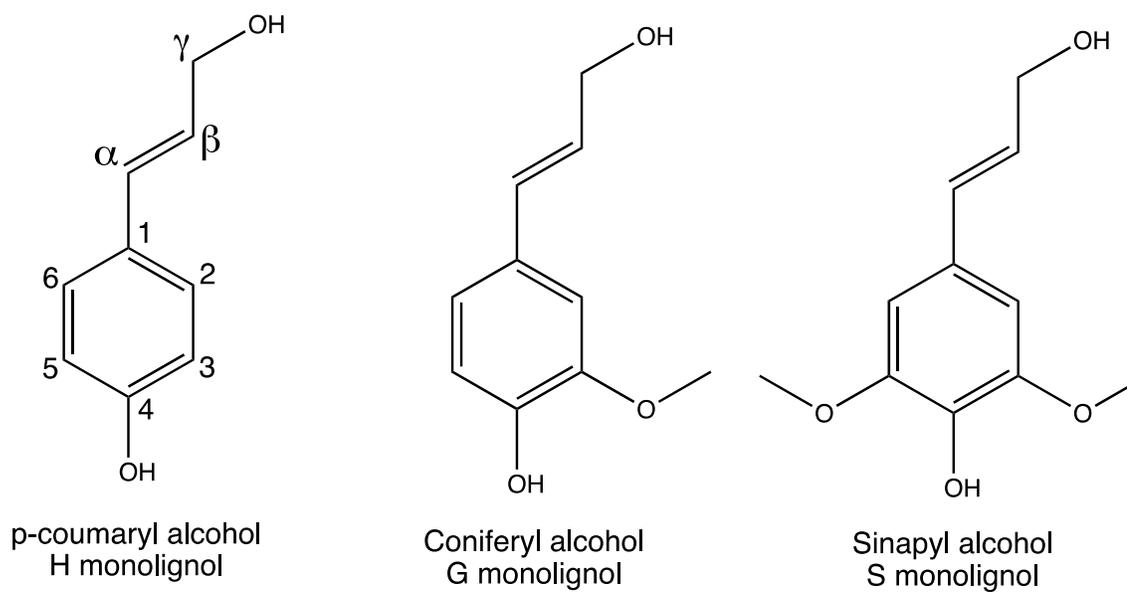
Figure 1.1 Structure of three monomeric units of lignin

p-coumaryl alcohol
H monolignol

Coniferyl alcohol
G monolignol

Sinapyl alcohol
S monolignol
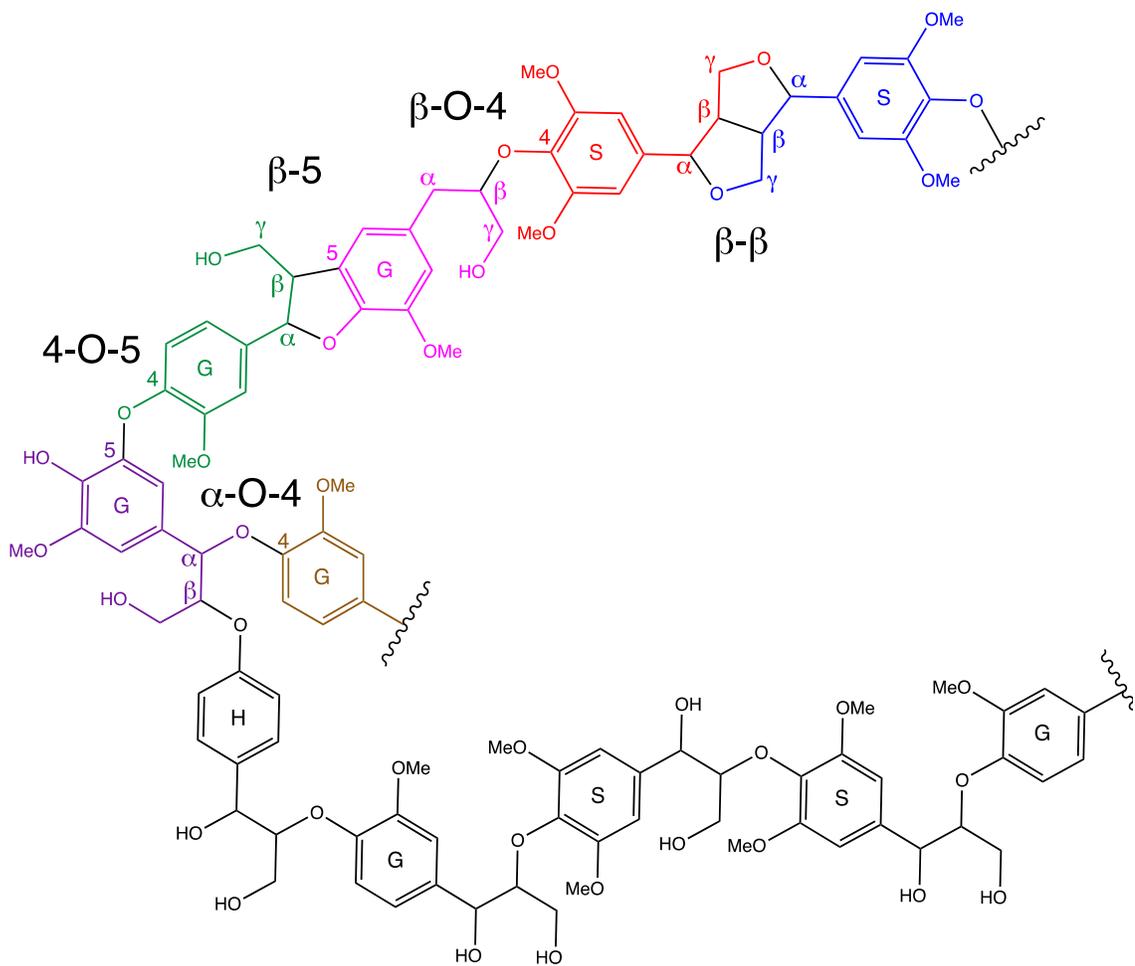
Figure 1.2 Simplified structure of lignin with labeled different bond types.

1.1.2 Pretreatment

As stated previously, lignocellulosic biomass has emerged as one of the most promising alternative carbon sources that are available in sufficient quantities and widely distributed[9]. A biorefinery can mediate the use of this abundant resource and move toward carbon-neutral energy and chemical production[28]. However, lignocellulosic biomass is a recalcitrant mixture of lignin, cellulose, and hemicellulose which are covalently bonded together. Therefore, a pretreatment step is required to deconstruct the biomass into its primary constituents[29].

Ideally, pretreatment is required to separate these three components into a structurally intact form to be utilized for various purposes. However, all pretreatment methods tend to alter the structure of the isolated components[30]. Historically, cellulose has been the focus of biomass utilization and most of the conventional pretreatments are cellulose-oriented, while lignin has been viewed as a waste byproduct[31]. One of the most popular cellulose-oriented pretreatments was pulping woody biomass to produce paper. In the year 2021, the global production of paper reached a staggering 415 million tons, resulting in a substantial amount of lignin as waste[32]. Given the urgent need to adopt sustainable practices, and the high value aromatic nature of lignin, there is an increasing emphasis on finding new applications for lignin, a highly abundant byproduct that has long been overlooked. As a result, there is a growing interest in the development of pretreatment methods with a lignin-first approach[9]. These methods not only enable the efficient production of higher quality lignin for various industrial applications but also allow for the recovery of other valuable components, such as cellulose and hemicellulose.

Organosolv pretreatment is gaining significant attention as a preferred approach for the pretreatment of biomass. Organosolv pretreatments are a group of biomass conversion methods that use organic solvents such as acids, alcohols, ketones, aldehydes, amines, ethers, or esters, etc., to fractionate the lignocellulosic biomass into its constituent components[33-37]. They typically involve three main steps: solvent impregnation, extraction, and recovery. In the solvent impregnation step, the biomass is saturated with the organic solvent, which can be a single solvent or a mixture of solvents. In the extraction step, the solvent-laden biomass is subjected to heat and/or pressure to break down the lignin-hemicellulose complex and release the sugars. Once the extraction is completed, the biomass residue can be recovered for further processing or disposal.

Hydrotropic pretreatment is a relatively new method for the pretreatment of lignocellulosic biomass, which uses a class of compounds called hydrotropes to solubilize the lignin and hemicellulose in biomass[38]. Hydrotropes are amphiphilic molecules that can form micelles in water, allowing them to interact with and dissolve the lignocellulosic components[39]. This method can be carried out at lower temperatures and atmospheric pressure, making it a more energy-efficient and environmentally friendly option compared to traditional pretreatment methods.

Acetic/formic acid pretreatment, developed by the Compagnie Industrielle de la Matière Végétale (CIMV) and referred to as CIMV pretreatment in some articles[40]. It involves the treatment of lignocellulosic biomass using an acetic acid/formic acid/water mixture. This method has been demonstrated to be highly efficient for various species of biomass, particularly annual plant wastes such as cereal straws, sugar cane, and sweet sorghum bagasses. This process has been shown to effectively break down the complex

structure of the biomass, leading to increased yields of fermentable sugars for subsequent downstream processes[40].

Lignin-first pretreatments involve fractionating biomass to extract lignin and other components while minimizing structural changes to lignin[41]. Various analytical techniques such as NMR, FTIR, and mass spectrometry are commonly used to evaluate the extracted lignin for structural analysis, while methods like GPC and MALDI mass spectrometry are employed for determining molecular weight distribution[21-24, 42]. In this dissertation, Chapter 4 presents a quantitative Derivatization Followed by Reductive Cleavage (qDFRC) method that was developed to assess lignin extracted using five different pretreatment methods. The results of this study will be discussed in greater detail in chapter four.
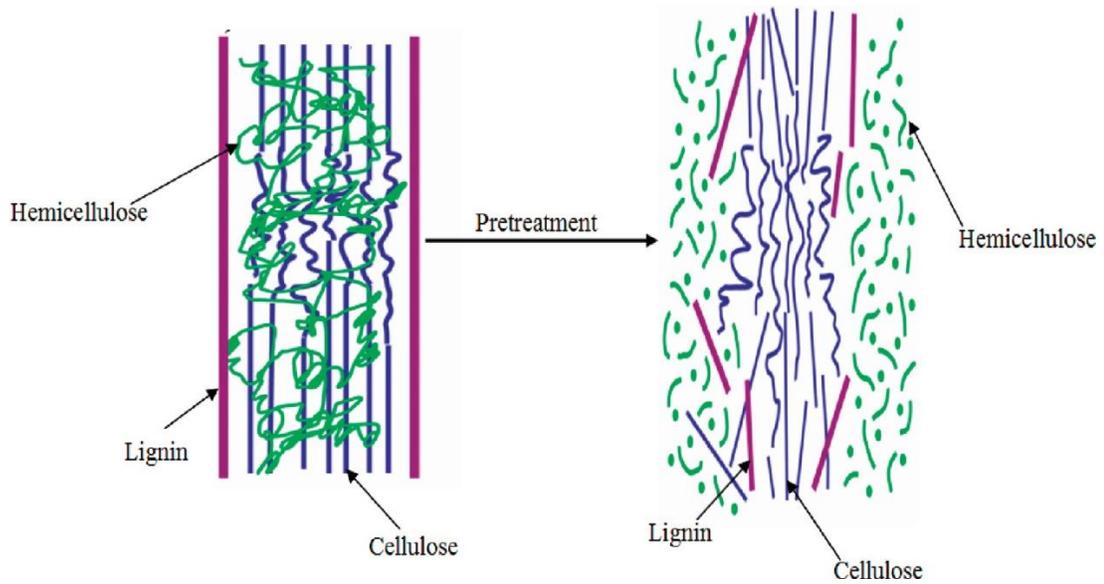
Figure 1.3 Various components of biomass and the effect of pretreatment[43]. "Reprinted with permission from Kumar et al. "Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production." *Industrial & engineering chemistry research* 48.8 (2009): 3713-3729.. Copyright 2009 American Chemical Society."

1.2   Mass Spectrometry

Mass spectrometry (MS) is an analytical technique that allows for the identification and analysis of chemical compounds based on their mass-to-charge ratio. The history of mass spectrometry dates back to the early 20th century when J.J. Thomson developed a device known as a mass spectrograph to study the behavior of ions in electric and magnetic fields. Later, in 1912, he developed the first true mass spectrometer, which allowed for the accurate determination of atomic masses and led to the discovery of numerous isotopes[44].

Today, mass spectrometry is an essential tool in many fields of science, including chemistry, biology, and medicine[45]. It is used for a wide range of applications, from drug development and environmental monitoring to forensic analysis and food safety testing. The technique has been instrumental in sequencing macromolecules and established itself as the primary technique in the proteomics field[46].

1.2.1   Main Components of a Mass Spectrometer

A mass spectrometer is an analytical instrument that is used to identify and quantify the chemical composition of a sample. The instrument typically consists of three main components: an ion source, a mass analyzer, and a detector. The ion source generates ions from the sample, and the mass analyzer separates the ions based on their mass-to-charge ratio (m/z) by subjecting them to an electric field. Finally, the detector detects and records the ions that pass through the mass analyzer. Depending on the specific design of the mass spectrometer, it may also include other components such as ion guides, collision cells, and ion traps to improve the sensitivity and specificity of the analysis. Mass spectrometry is a

powerful tool for a wide range of applications, including proteomics, metabolomics, drug discovery, and lignin structural analysis.

1.2.1.1   Ion Source

The ion source is a critical component of a mass spectrometer, responsible for generating ions from the sample to be analyzed. There are two main categories of ionization: soft ionization and hard ionization[47]. Soft ionization techniques, such as electrospray ionization (ESI), are typically used for the analysis of large biomolecules such as proteins, peptides, and nucleic acids, while hard ionization techniques, such as electron ionization (EI), are typically coupled with gas chromatography and used for the analysis of small molecules and volatile compounds.

Electron Ionization (EI) is a hard ionization technique that is commonly used for the analysis of small molecules. In EI, the gaseous sample passes through the ion source chamber where a high voltage (70 eV) is applied to a filament, resulting in a high-energy electron beam. The electron beam and the sample effluent are perpendicular, and at the intersection of these two flows, an electron is ejected from the analyte, forming radical cations[48]. It is noteworthy that a magnet is used to cause the electron beam to have a spiral-like trajectory, increasing contact with the sample. The generated ions are focused and ejected toward the mass analyzer with an electric potential applied by the repeller plate at a 90° angle[49].

Electron Ionization predominantly generates singly charged ions. Due to its instability, the molecular ion is often not observed at high intensity, or not observed at all.

Typically, the unstable molecular radical cations undergo one of three pathways: 1) loss of neutral radical fragments, 2) elimination of a charged radical, or 3) intermolecular rearrangements[48]. The applied vacuum and the short lifetime of the molecular radical cation eliminate any matrix effect on the analyte, while the unique bond energies in the analyte produce distinct fragmentation patterns that can provide fingerprint structural information[48]. In this dissertation, EI ionization is used with gas chromatography to characterize lignin model compound precursors and will be discussed in Chapter 2.

ESI is a soft ionization technique that is widely used for the analysis of large and complex molecules, such as proteins and nucleic acids. During ESI, ionization occurs in solution through the donation or acceptance of protons, resulting in the formation of protonated or deprotonated ions. Alternatively, ions can be obtained by binding the analyte with ions such as Li^+, K^+, or Na^+ to form adducts. The analyte solution is then introduced into the mass spectrometer via a capillary. A high voltage applied to the capillary tip forms a Taylor cone in atmospheric pressure which has a high charge density in the tip of the cone and results in the ejection of a charged solvent droplet, the droplets shrink rapidly and subsequently generate gaseous ions[50, 51]. The molecular ion is often observed in this type of ionization. Unlike EI, the efficiency of ESI ionization is strongly influenced by the structure of the analyte, and the type of ion formation, which can lead to biased abundance for certain analytes[52]. In this dissertation, ESI ionization was employed with a Q-Exactive orbitrap mass spectrometer to characterize and structural analysis of a synthetic lignin oligomer and will be discussed in more detail in Chapter 3.

The choice of ionization technique depends on the nature of the sample being analyzed, as well as the specific requirements of the analysis. By using different ionization techniques, mass spectrometry can be applied to a wide range of samples, from small molecules to large biomolecules.
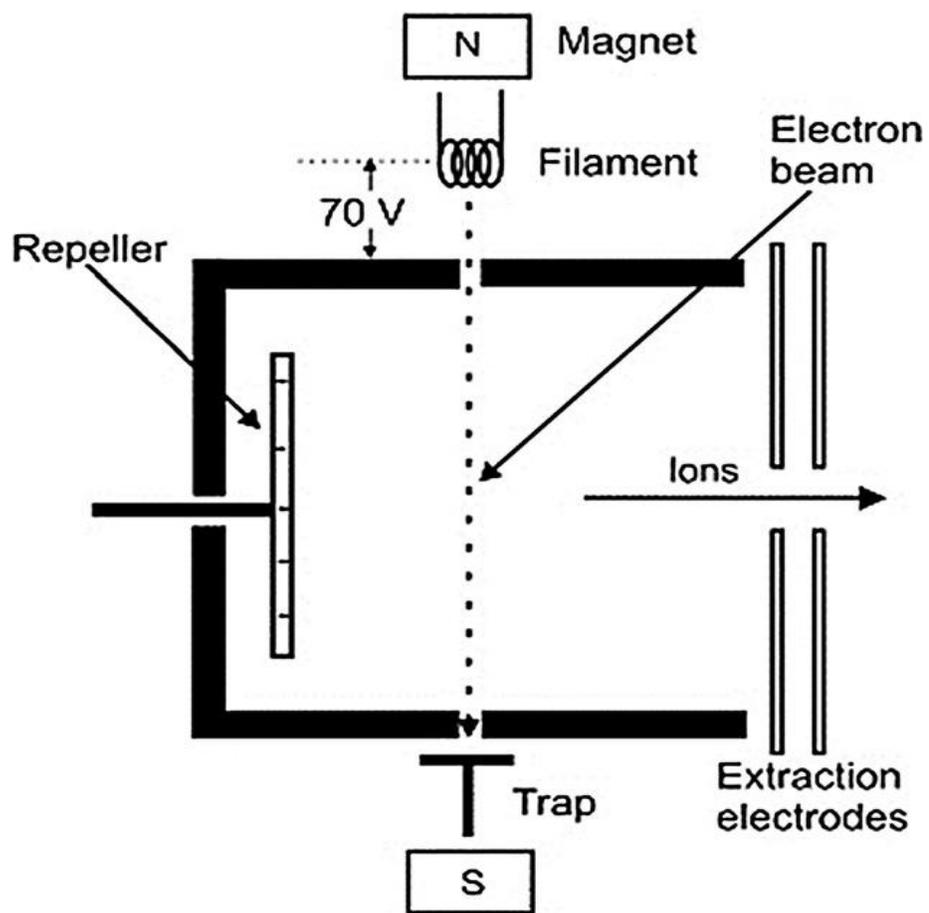
Figure 1.4 Schematic illustration of ionization chamber in Electron Ionization (EI)[49].
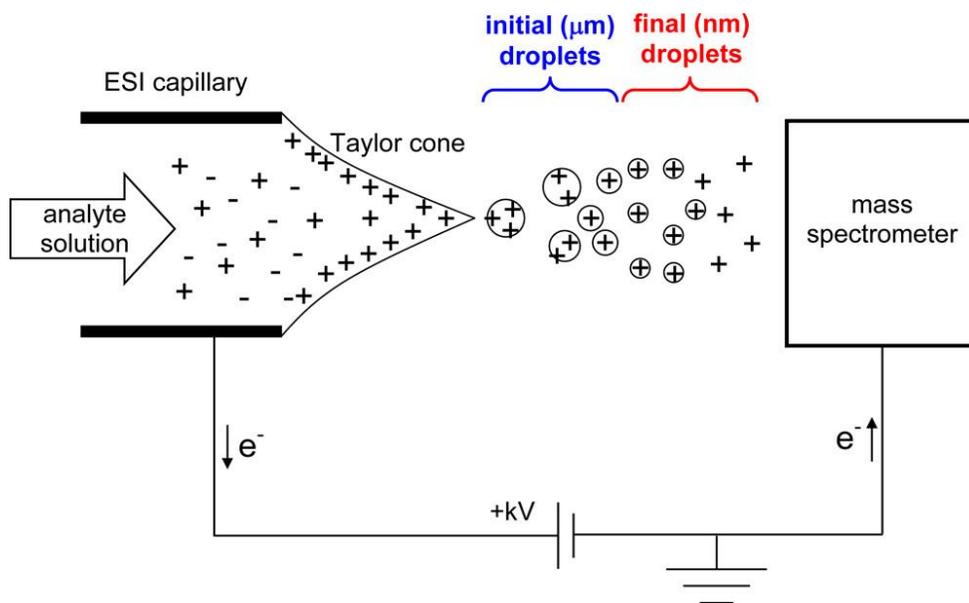
Figure 1.5 Schematic illustration of ESI ionization source[53]. "Reprinted with permission from Konermann et al. "Unraveling the mechanism of electrospray ionization." (2013): 2-9.Copyright 2013 American Chemical Society."

1.2.1.2   Mass Analyzer

The mass analyzer is definitely the most important component of a mass spectrometer. It separates ions of different masses based on their mass-to-charge ratio (m/z). The development of different types of mass analyzers has allowed for a wide range of applications of mass spectrometry, from small molecule analysis to proteomics and metabolomics.

Different types of mass analyzers have varying levels of sensitivity, resolution, and accuracy, and they are the main contributor to the cost of a mass spectrometer. For example, quadrupole mass analyzers are relatively simple and inexpensive, and they are widely used in many applications. On the other hand, Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometers are among the most expensive instruments, but they offer exceptional mass resolution and accuracy, making them invaluable for a wide range of applications, including proteomics and metabolomics[54].

The choice of the mass analyzer for a specific application depends on the analytical goals. While quadrupole mass analyzers are commonly used in routine analysis, high-end applications may require more sophisticated and expensive mass analyzers, such as FT-ICR. For the purpose of this dissertation, a quadrupole mass analyzer coupled with gas chromatography is used for characterizing lower molecular mass synthesized compounds, and high resolution orbitrap mass analyzer in Q-Exactive orbitrap mass spectrometer (Figure 1.6) is used for higher molecular accurate mass mainly in chapter 3.

A quadrupole is a type of mass analyzer commonly used in mass spectrometry. It consists of four cylindrical rods arranged parallel to each other and perpendicular to the

direction of ion movement. The rods are separated by a small gap through which the ions pass. Radiofrequency (RF) and DC voltage are applied to the rods in such a way that the ions are selectively filtered based on their mass-to-charge ratio. Only ions with a certain mass-to-charge ratio pass through the quadrupole and reach the detector, while other ions are deflected and do not reach the detector[55, 56]. This allows for the separation and detection of ions based on their mass-to-charge ratio, with selectivity for one mass-to-charge ratio unit.

The quadrupole mass analyzer can be used either as a standalone unit or in combination with other mass analyzers such as ion traps or multiple quadrupoles (triple Quad) to provide higher resolution and the ability to perform tandem mass spectrometry. Tandem mass spectrometry involves using two or more mass analyzers in succession to fragment and analyzes ions, which can provide more detailed information about the structure of a molecule. When used in combination with other mass analyzers, the quadrupole can either focus the ion flow or isolate a desired mass-to-charge ratio for analysis.

The Orbitrap mass analyzer is a relatively new type of mass analyzer that has gained popularity in recent years due to its high resolution and mass accuracy. It consists of three electrodes: two outer electrodes shaped like cups facing each other and a spindle-like central electrode that holds the trap together. Applying a voltage between the outer and central electrodes generates a strictly linear electric field along the axis, resulting in rotational trajectory of ions circulating around the central electrode. In addition, the conical shape of the electrodes produces an axial electric field that directs ions toward the widest region of the trap and initiates harmonic axial oscillations. The oscillations of the ions

produce frequencies corresponding to their m/z. By detecting the frequencies of these oscillations, the mass-to-charge ratios of the ions can be determined using Fourier transform[57-59].

The Orbitrap mass analyzer offers several advantages over other mass analyzers. One of its major strengths is its ability to provide high resolution accurate mass (HRAM) data, which is a powerful tool for structural analysis of analytes. HRAM data is particularly useful in identifying unknown compounds, determining the elemental composition of analytes, and detecting small variations in molecular weight.

In this dissertation, Q-Exactive Orbitrap mass spectrometer was used for structural analysis of a lignin trimer and discussed in Chapter 3
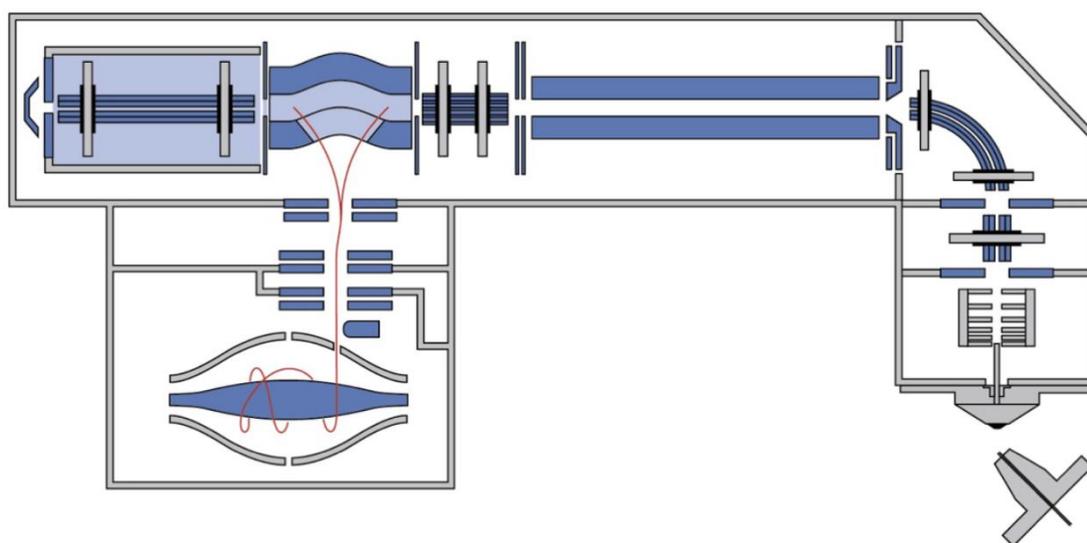
Figure 1.6 Schematic diagram of Thermo Fisher Q-Exactive mass spectrometer[60]. Adapted with permission from Hecht et al. "Fundamentals and advances of orbitrap mass spectrometry." *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation* (2006): 1-40. Copyright © 2019 John Wiley & Sons, Ltd.

CHAPTER 2. CHARACTERIZATION OF LIGNIN MODEL COMPOUND
PRECURSORS USING GAS CHROMATOGRAPHY-MASS SPECTROMETRY

2.1   Introduction

Currently, the primary source of aromatic compounds comes from finite fossil fuels, and there is a growing need to replace these finite sources with renewable ones. Lignin, as the most abundant aromatic polymer in nature, has the potential to serve as a valuable source of aromatic chemicals and pharmaceutical synthons[1]. Lignin is one of the three main components of biomass, next to cellulose and hemicellulose. It consists of three main monomers as p-hydroxy coumaryl, coniferyl, and sinapyl alcohol, also known as H, G and S monolignols, respectively[61]. These monomeric units are link together with various bond type, with the β-O-4 type being the most abundant one. Hence, investigating the β-O-4 linkage is crucial for the development of structural analysis and enhancing our understanding of lignin.

The structural analysis of lignin and lignin breakdown products has been significantly aided by the use of lignin model compounds. Ideally, a model compound should mimic the structure of natural lignin to enable accurate analysis[1, 62-64]. Numerous model compounds have been developed and reported; however, many of these compounds lack the double bond found in the aliphatic chain of natural lignin[65-67]. The double bond in the aliphatic chain is a key factor that plays a significant role in the biosynthesis of lignin. Therefore, it is of great importance that model compounds used for structural analysis of lignin should contain this characteristic double bond.

20

2.1.1 Aldol Addition Reaction

The aldol reaction is a carbon-carbon bond-forming reaction that involves the addition of an enol or an enolate to a carbonyl compound in aldehydes or ketones. The name "aldol" comes from the combination of the words "aldehyde" and "alcohol," which are the two functional groups involved in this reaction.

The general mechanism of the aldol reaction involves the formation of an enolate or enol intermediate, which attacks the carbonyl group of another molecule to form a β-hydroxyaldehyde or β-hydroxyketone. The reaction can be catalyzed by a base to remove a hydrogen from alpha carbon and generate an enolate[64, 68].

The aldol reaction can also proceed through a crossed aldol reaction, in which two different carbonyl compounds are used. This allows for the creation of more complex structures.

The current study aimed to synthesize precursors that can be utilized in aldol addition. Specifically, the synthesized precursor contains an alpha hydrogen that is adjacent to an ester carbonyl, allowing for the formation of an enolate intermediate. This enolate subsequently attacks the carbonyl of an aldehyde to form a β-hydroxy compound, which bears structural similarity to the β-O-4 bond type observed in lignin. These precursors were synthesized with the intention of enabling the synthesis of β-O-4 model oligomer sequences to further enhance structural analysis of lignin-related compounds.

a

b

Aldol reaction

Aliphatic-terminus precursor

Phenol-terminus precurs

alpha hydrogen

Aldol reaction

β-O-4 bond type
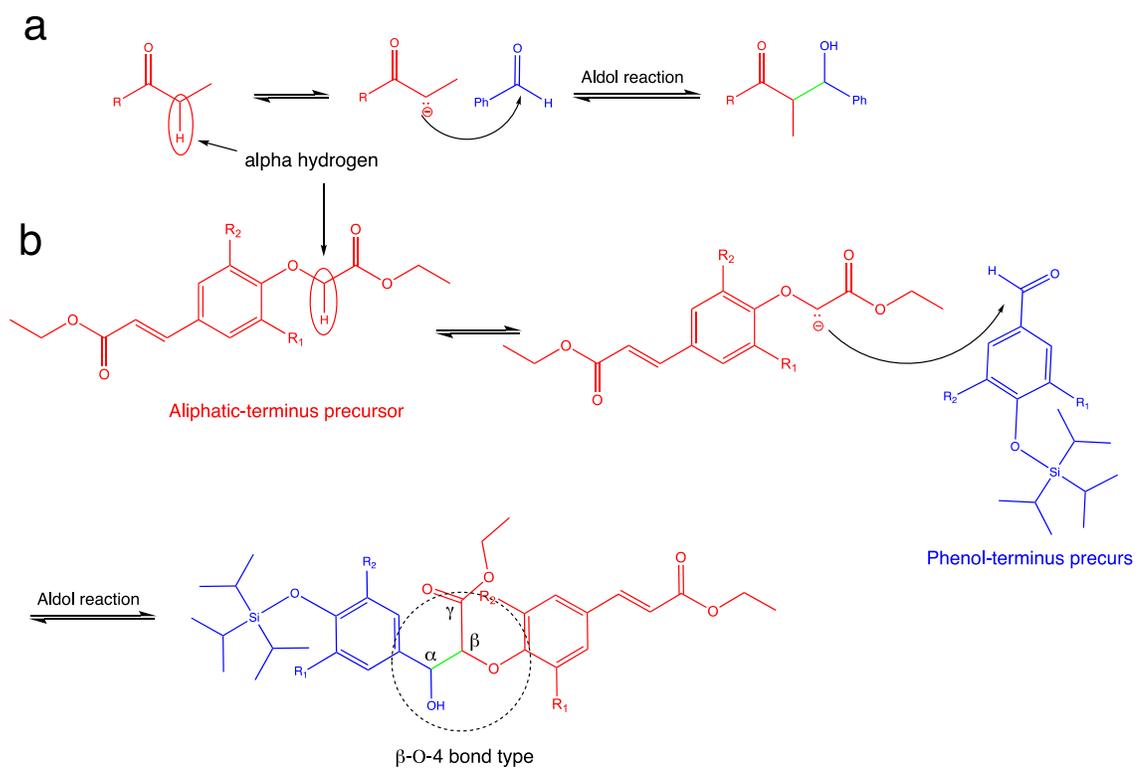
Figure 2.1 General schematic of base catalyzed cross aldol reaction(a). Cross aldol reaction using synthesized precursors (b).

2.1.2 Nomenclature of Precursors

Synthesis of β-O-4 model compounds carried out using aldol coupling reaction. To accomplish the coupling reaction, precursors were made in three different categories. The structure of a β-O-4 oligomer can be described as a chain of monomeric units linked together by β-O-4 bonds. The oligomer chain contains a phenol group at one end and an aliphatic propenol chain at the other end. Therefore, the monomeric unit on the phenol end of the oligomer was named phenol-terminus, and the monomeric unit on the aliphatic chain end was named aliphatic-terminus. The monomeric units in the middle of the chain are repeating units that form a β-O-4 bond with the phenol functional group on the phenol-terminus side and a β-O-4 bond with the aliphatic chain on the aliphatic-terminus side and they were named as the middle ring.

These types of precursors were synthesized for each type of monomeric units described above. A β-O-4 dimer model compound can be synthesized by coupling of a phenol-terminus precursor with an aliphatic-terminus precursor, while synthesis of a trimer β-O-4 model compound requires coupling of a phenol-terminus precursor with a middle ring precursor and then the synthesized dimeric unit can be coupled with an aliphatic-terminus precursor. Hypothetically, the middle ring can be a repeating unit to synthesize larger oligomers.

Phenol-terminus   Repeating middle ring   Aliphatic-terminus
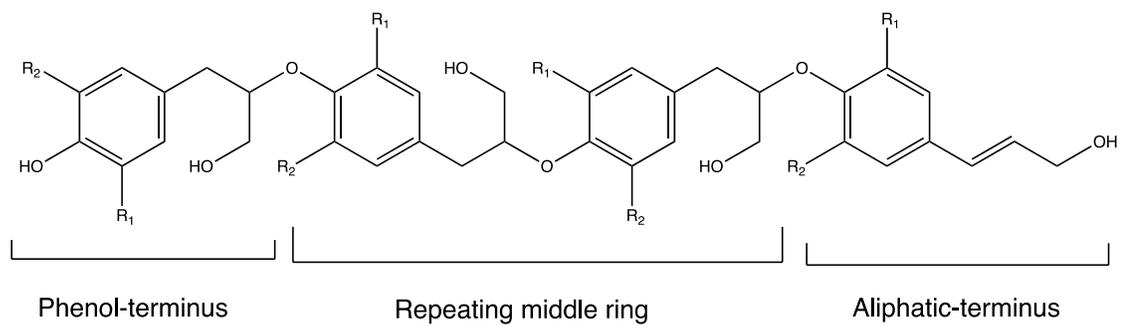
H: R1, R2= H
G: R_1= OMe, R_2= H
S: R_1, R_2= OMe

Figure 2.2 Hypothetical β-O-4 polymer chain and nomenclature of the precursors

Oligomer:

Phenol-teminus precursor   Repeating middle ring precursor   Aliphatic-terminus precursor

Dimer:

H: R1, R2= H
G: R_1= OMe, R_2= H
S: R_1, R_2= OMe

Phenol-teminus precursor   Aliphatic-terminus precursor
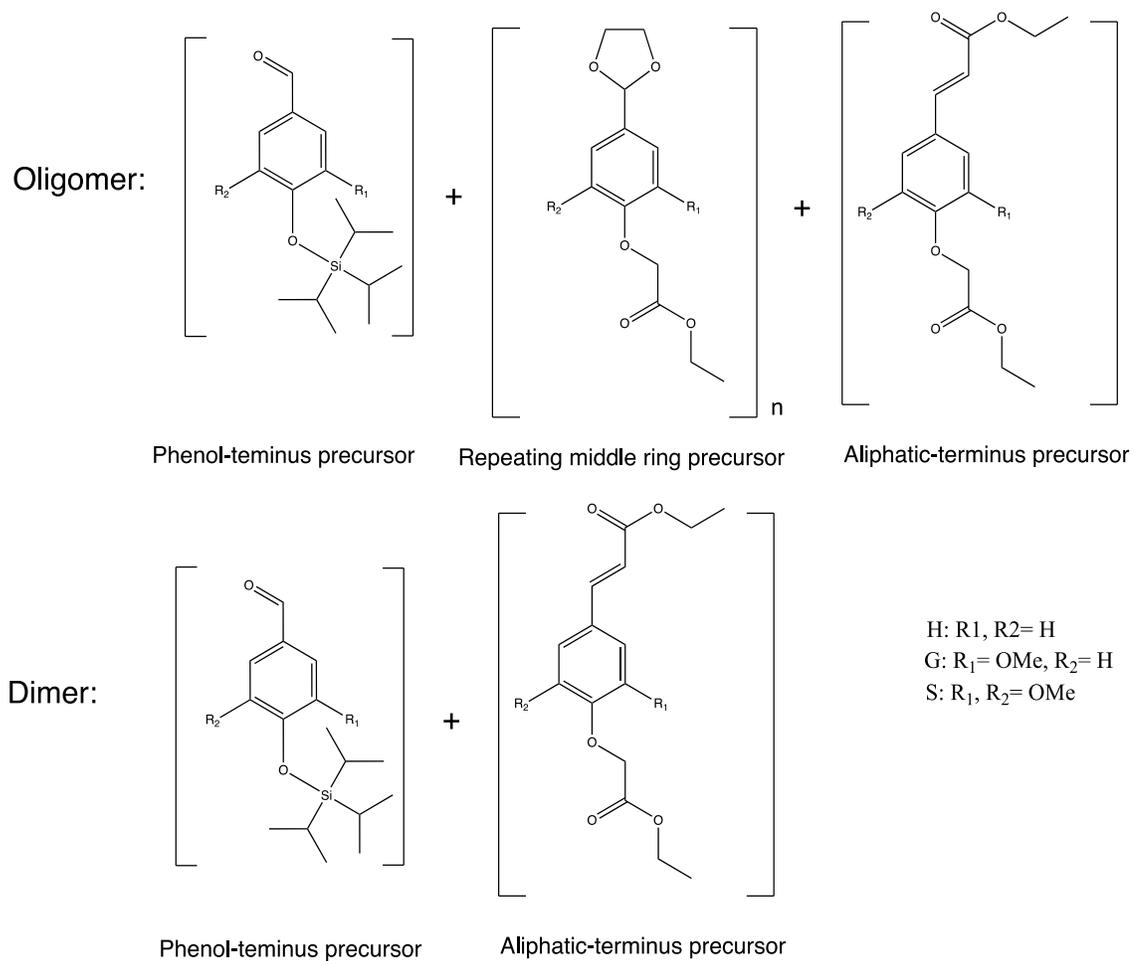
Figure 2.3 General order of addition of precursors to synthesize β-O-4 dimers and oligomers

2.2   Material and Methods

2.2.1   Chemicals

All chemicals were used without further purification. Vanillin, 4-hydroxy benzaldehyde, ferulic acid, ammonium chloride, ethanol, potassium carbonate, ethyl bromoacetate, and triethylamine (TEA) were purchased from Alfa Aesar (Ward Hill, MA, USA). Sinapic acid, imidazole, ethyl acetate, acetyl bromide, acetone, toluene, ethylene glycol, p-toluenesulfonic acid monohydrate (p-TsOH), phenylboronic acid, cupric acetate, pyridine, and sodium sulfate were purchased from Fisher Scientific (Fair Lawn, NJ, USA). Benzyl bromide, p-coumaric acid, diisopropylethylamine (DIPEA), and 4-dimethylaminopyridine (DMAP) were purchased from Sigma-Aldrich (Saint Louis, MO, USA). Methylene chloride, and triisopropylsilyl chloride (TIPSCl) were purchased from VWR chemicals (Radnor, PA, USA) and syringe aldehyde was purchased from Dofine Chem (Shanghai, China).

2.2.2   Methods

2.2.2.1   Synthesis of Phenol-terminus Precursor

Corresponding aldehyde (1 mol) was dissolved in methylene chloride (1:20 w/v). 4-hydroxy benzaldehyde, vanillin, and syringaldehyde were used for H, G, and S precursors, respectively. Imidazole (1.5 mol) and triisopropylsilyl chloride (TIPSCl)(4 mol) were added to the reaction mixture. The solution was stirred overnight and monitored

with TLC plate. The product was extracted with ethyl acetate (100 mL × 3 times) and saturated ammonium chloride (50 mL) and purified with a silica gel gravimetry column.

H unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{16}H_{26}O_2Si]^{\bullet+}$ m/z 278, obs. mass m/z 278.

G unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{17}H_{28}O_3Si]^{\bullet+}$ m/z 308 (not observed), $[M-43]^+$ obs. mass m/z 265.

S unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{18}H_{30}O_4Si]^{\bullet+}$ m/z 338 (not observed), $[M-43]^+$ obs. mass m/z 295.

H: R1, R2= H
G: R_1= OMe, R_2= H
S: R_1, R_2= OMe
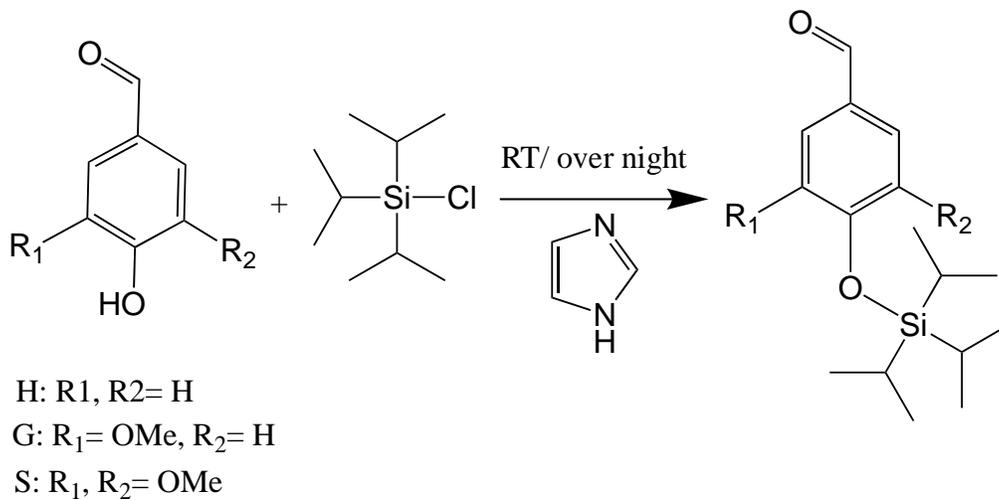
Figure 2.4 Synthesis of phenol terminus precursor.

2.2.2.2   Synthesis of Aliphatic-terminus Precursor

To synthesize aliphatic-terminus precursors, p-coumaric acid, ferulic acid, and sinapic acid were used for H, G, and S units, respectively. Synthesis of the precursor was accomplished by two steps: 1) conversion of acid functional group to ethyl ester 2) addition of an ester group to the phenol functional.

2.2.2.2.1   STEP 1

The corresponding acid (5 g) was dissolved in ethanol (50 mL) (1:10 v/w ratio). Ethanol was heated to promote dissolution of the solute. After complete dissolution, the solution was cooled down to room temperature, then acetyl bromide (2.5 mL) was added dropwise. The solution was stirred overnight and monitored by TLC. After completion, the solution was concentrated under reduced pressure to approximately 20 mL to achieve phase separation in extraction step. The product was extracted with ethyl acetate (50 mL × 3 times) and potassium carbonate (50 mL). The product was used for the next step without any purification.

H unit: GC-MS $[M]^{\bullet+}$ calc. mass for TMS derivatized compound $[C_{14}H_{20}O_3Si]^{\bullet+}$ m/z 264, obs. mass m/z 264.

G unit: GC-MS $[M]^{\bullet+}$ calc. mass for TMS derivatized compound $[C_{15}H_{22}O_4Si]^{\bullet+}$ m/z 294, obs. mass m/z 294.

S unit: GC-MS $[M]^{\bullet+}$ calc. mass for TMS derivatized compound $[C_{16}H_{24}O_5Si]^{\bullet+}$ m/z 324, obs. mass m/z 324.

2.2.2.2.2 STEP 2

The product from the step 1 (1 mol) was dissolved in acetone with the 1:20 w/v ratio. Ethyl bromoacetate (1.5 mol) and potassium carbonate (1.5 mol) were added. The solution was on reflux for 4 h and monitored with microextraction and GC-MS. The product was extracted with methylene chloride (100 mL × 3 times) and saturated ammonium chloride (100 mL) and purified with silica gel gravity column chromatography (yield 92%).

H unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{15}H_{18}O_5]^{\bullet+}$ m/z 278, obs. mass m/z 278.

G unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{16}H_{20}O_6]^{\bullet+}$ m/z 308, obs. mass m/z 308.

S unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{17}H_{22}O_7]^{\bullet+}$ m/z 338, obs. mass m/z 338.
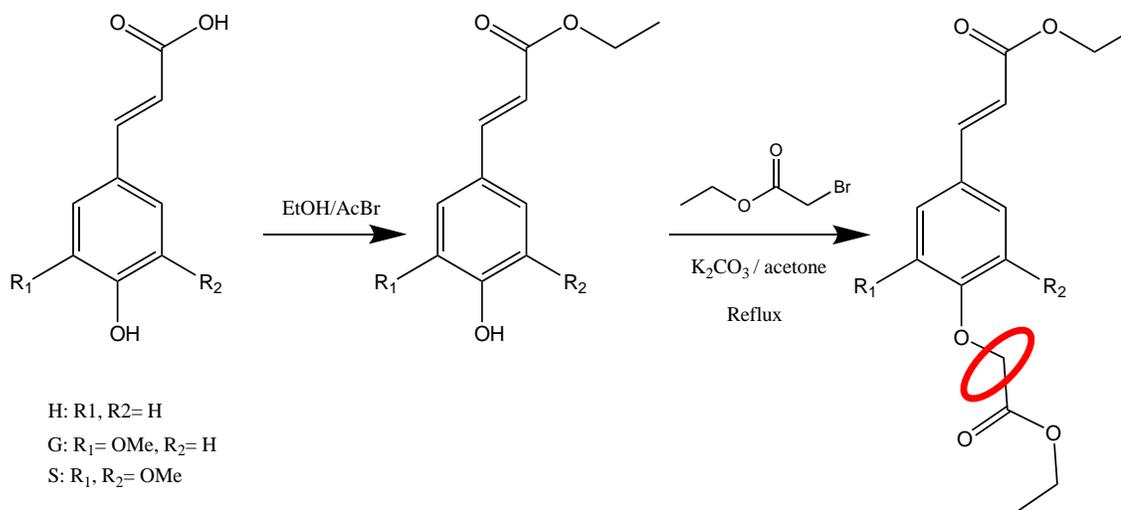
Figure 2.5 Synthesis of aliphatic-terminus precursor. The highlighted carbon is alpha carbon required for aldol reaction.

2.2.2.3   Synthesis of Middle Ring Precursor

To synthesize middle ring precursors, 4-hydroxy benzaldehyde, vanillin, and syringe aldehyde were used for H, G and S units, respectively. The middle ring synthesis was accomplished in two steps: 1) addition of ester to the phenol functional group. 2) protection of the aldehyde functional group with ethylene glycol.

2.2.2.3.1   STEP 1

The starting corresponding aldehyde (4-hydroxy benzaldehyde, vanillin, or syringe aldehyde) (1 mol) was dissolved in acetone (1:10 w/v). Ethyl bromoacetate (1.5 mol) and potassium carbonate (1.5 mol) were added to the solution. The solution was on reflux for three hours and monitored with TLC plate. The mixture was cooled down to room temperature and was filtered (Fisherbrand filter paper, qualitative: P8, porosity coarse flow rate: fast, Pittsburg, PA, USA) with gravity filtration to remove potassium carbonate and rinsed with ethyl acetate. The solvent was removed under reduced pressure. Subsequently, a small quantity of ethanol was added to the resulting residue and was kept in -20°C overnight to form a white solid. The solid product was filtered and rinsed with cold ethanol and air-dried (83% yield).

H unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{11}H_{12}O_4]^{\bullet+}$ m/z 208, obs. mass m/z 208.

G unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{12}H_{14}O_5]^{\bullet+}$ m/z 238, obs. mass m/z 238.

S unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{13}H_{16}O_6]^{\bullet+}$ m/z 268, obs. mass m/z 268.

The product from step 1 (1 mol) was dissolved in toluene (1:25 w/v). Ethylene glycol (10 mol) and p-toluenesulfonic acid monohydrate (0.05%wt) was added. The solution was on reflux using a Dean-Stark apparatus for three hours and monitored with TLC. The reaction was quenched with potassium carbonate (5 mL, 100 mg/mL) and then extracted with ethyl acetate (100 mL × 3 times) and ammonium chloride (100 mL). The collective organic solvent was removed under reduced pressure, and a small quantity of ethanol was added to the resulting residue and kept at -20°C overnight to form a solid (85% yield).

H unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{13}H_{16}O_5]^{\bullet+}$ m/z 252 (not observed), $[M-1]^+$ obs. mass m/z 251.

G unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{14}H_{18}O_6]^{\bullet+}$ m/z 282 (not observed), $[M-1]^+$ obs. mass m/z 281.

S unit: GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{15}H_{20}O_7]^{\bullet+}$ m/z 312, obs. mass m/z 312.

H: R1, R2= H
G: R$_1$= OMe, R$_2$= H
S: R$_1$, R$_2$= OMe
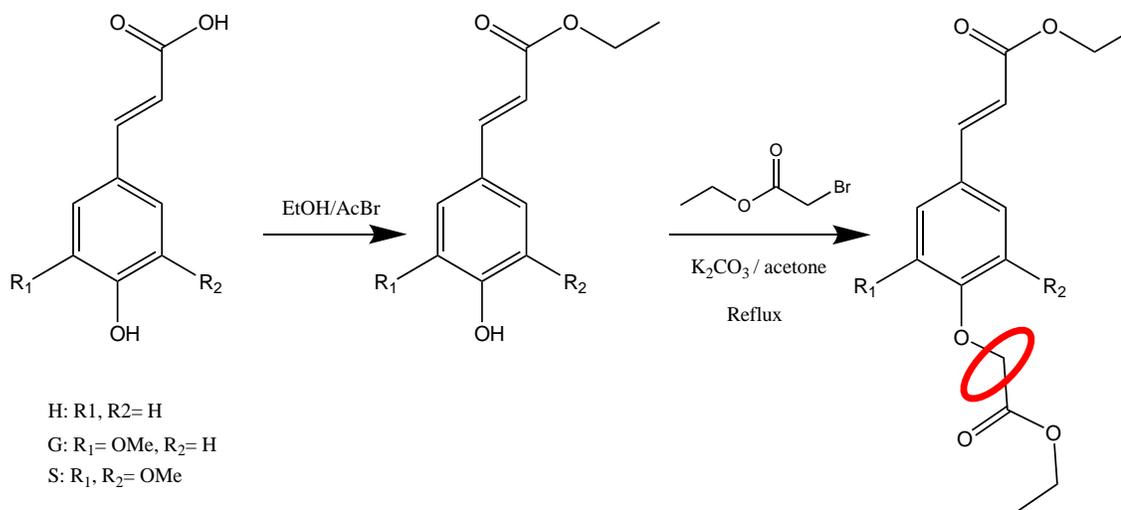
Figure 2.6 Synthesis of middle ring precursor. The highlighted carbon is the alpha carbon required for aldol reaction.

2.2.2.4 Synthesis of Benzyl-G-βO4-G Dimer

The G-βO4-G model compound was synthesized by ferric chloride coupling of G monolignol as described in Section 3.2.2.2.3. The G-βO4-G dimer model compound (200 mg, 0.56 mmol) was dissolved in acetone (10 mL). Benzyl bromide (99.4 μL, 0.84 mmol) and potassium carbonate (116 mg, 0.84 mmol) were added to the solution. The reaction mixture was on reflux for 24 hours under nitrogen and monitored with TLC. The product was extracted with ethyl acetate (15 mL × 3 times) and saturated ammonium chloride (10 mL). The product was purified with silica gel gravity column (89% yield).

GC-MS $[M]^{•+}$ calc. mass for TMS derivatized compound $[C_{36}H_{54}O_7Si_3]^{•+}$ m/z 682, obs. mass m/z 682.

Q-Exactive (ESI-Orbitrap) $[M+Li]^+$ calc. for $[C_{27}H_{30}O_7Li]^+$ m/z 473.2146, obs. mass m/z 473.2149. Error (ppm)= 0.6465

2.2.2.5 Synthesis of Phenyl-G-βO4-G

Phenyl-G-βO4-G dimer was prepared by Chan-Lam coupling of G-βO4-G dimer with phenylboronic acid as reported by Evans et al.(1998). Briefly, model G-βO4-G dimer (73 mg, 0.19 mmol) was dissolved in methylene chloride (2 mL). Phenyl boronic acid (47.4 mg, 0.39 mmol), cupric acetate (35.3 mg, 0.19 mmol), pyridine (78 μL, 0.97 mmol), and 4° A molecular sieve (2.4 mg) were added to the solution. The solution was stirred at room temperature for 18 h to form the product. The reaction was monitored with TLC. The product was purified with silica gel gravity column chromatography. (52% yield).

GC-MS $[M]^{•+}$ calc. mass for TMS derivatized compound $[C_{35}H_{52}O_7Si_3]^{•+}$ m/z 668, obs. mass m/z 668.
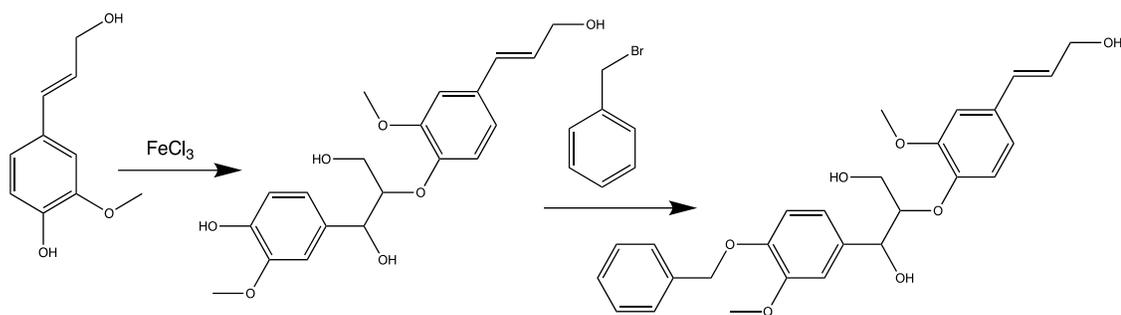
Figure 2.7 Synthetic pathway of benzyl-modified G-βO4-G dimer.
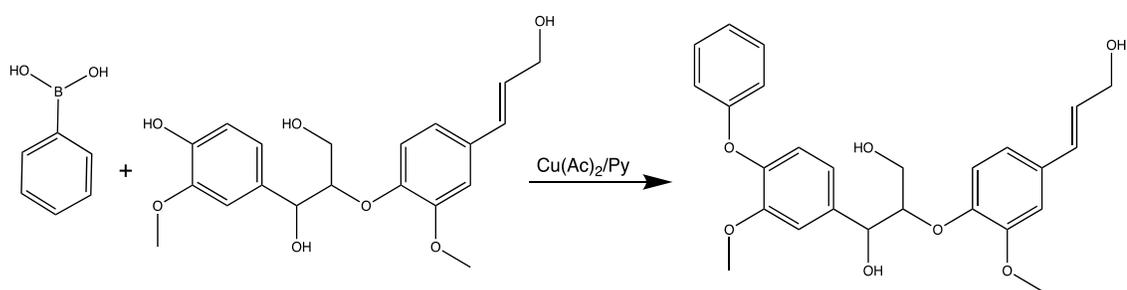
Figure 2.8 Synthesis of phenyl-modified G-βO4-G dimer with Chan-Lam coupling.

2.2.3   Gass Chromatography-Mass Spectrometry

The GC-MS analysis was conducted on an Agilent 5973 MSD equipped with HP 6890 GC and HP 7683 injector controlled by ChemStation D.03.00611. A DB-5HT (Agilent Santa Clara, CA, USA). A GC column with 15m length, internal diameter of 250 µm and film thickness of 0.1µm was used. The injector temperature was set to 250 °C. The oven temperature was set at 100 °C and held for 3 minutes, then ramped at 15 °C/min to 280 °C and held for 10 minutes for a total method time of 25 minutes. An injection split ratio of 50:1 was used for analysis.

2.3   Discussion

Model compounds have played a crucial role in the structural elucidation of lignin and its degradation products. The use of model compounds allows for a more controlled and systematic approach to studying the behavior of lignin. As analytical techniques continue to advance, model compounds will remain an essential tool in the study of lignin. Hence, there is a critical need to develop advanced model compounds that can effectively mimic the structure of natural lignin. One crucial aspect of the lignin structure is the presence of a double bond on the aliphatic chain. This bond plays a significant role in the biosynthesis of lignin by forming a resonance structure with the aromatic ring, leading to the formation of various bond types in lignin[14]. This characteristic feature, however, has been largely neglected in most of the reported model compounds in the literature. In this study, various precursors for the H, G, and S units, which can be employed to synthesize a model compound of the β-O-4 type, have been synthesized. The utilization of these precursors under controlled conditions through aldol reaction allows for the synthesis of a

desired sequence of a β-O-4 oligomer, which significantly contributes to the structural elucidation of lignin[62, 69-71].

The aldol reaction that can be employed to synthesize a β-O-4 bond type, requires a carbonyl and alpha carbon site to an ester. The carbonyl group was provided in phenol-terminus by 4-hydroxy benzaldehyde, vanillin, and syringe aldehyde for H, G, and S units. These precursors mimic lignin monomeric units with methoxy groups on the ring and the presence of a phenol group, and also provide the carbonyl by the aldehyde functional group. To prevent the phenol group from interfering in the coupling reaction, it was protected with a triisopropylsilyl (TIPS) group.

In the process of synthesizing the phenol-terminus precursor, an interesting byproduct was observed. The characterization of the byproduct was accomplished with various analytical techniques and discussed in detail.

The other precursor for the aldol reaction requires an alpha carbon to a carbonyl group. To mimic the monomeric units of lignin, p-coumaric acid, ferulic acid, and sinapic acid were employed as the initial reagent in the synthesis of aliphatic-terminus precursor. The structure of the initial acid compounds is similar to that of monolignols, except that they contain an acidic group rather than an aliphatic alcohol. It is noteworthy that monolignols can serve as initial reagent for the synthesis of aliphatic-terminus precursor. However, owing to the higher costs of monolignols, they were synthesized in-house using the aforementioned acid compounds, as described in 3.2.2.2.2. Consequently, ethyl esters of the acids were employed in the coupling reaction. The addition of an ethyl acetate group to the phenol functional group provided the alpha carbon necessary for the aldol reaction.

Furthermore, in this study, two pre-synthesized G-βO4-G dimers were subjected to structural modifications by the request of our collaborator in chemical engineering department. Specifically, a benzyl and a phenyl group were added to the phenol side of the dimer. The characterization of these dimers with GC-MS electron ionization revealed interesting structural information which will discussed later in this section.

2.3.1   Synthesis of Phenol-terminus Precursor

Three different precursors for H, G, and S unit were synthesized following the reaction presented in Figure 2.5. The product was analyzed with GC-MC. Figure 2.9 is the mass spectrum of the H unit of phenol-terminus precursor. The molecular ion mass calculated as m/z 278 and observed in GC-MS as m/z 278. The fragmentation pattern of analytes containing aldehyde functional groups commonly involves the loss of a radical hydrogen[72-75], resulting in the formation of [M-1]$^+$ ion. Notably, this fragmentation pattern was not observed for the H unit and the molecular ion appeared as m/z 278. Conversely, in the case of G and S units, the molecular ion was not detected; instead, [M-1]$^+$ ion and the characteristic loss of alkyl groups for silyl-protected analytes were observed. The H unit also showed characteristic loss of alkyl group from TIPS protecting group. Triisopropylsilyl (TIPS) group consists of three isopropyl groups. By fragmentation of one isopropyl from the analyte, m/z 235 appeared in the mass spectrum. It further fragmented by losing methyl groups and provided m/z 207 and m/z 179 as presented in Figure 2.9. The fragments at m/z 165 and m/z 149 were also the loss of methyl group followed by a hydrogen rearrangement. Similar results were observed for G and S units, with a m/z 30 difference attributed to the introduction of a methoxy group.
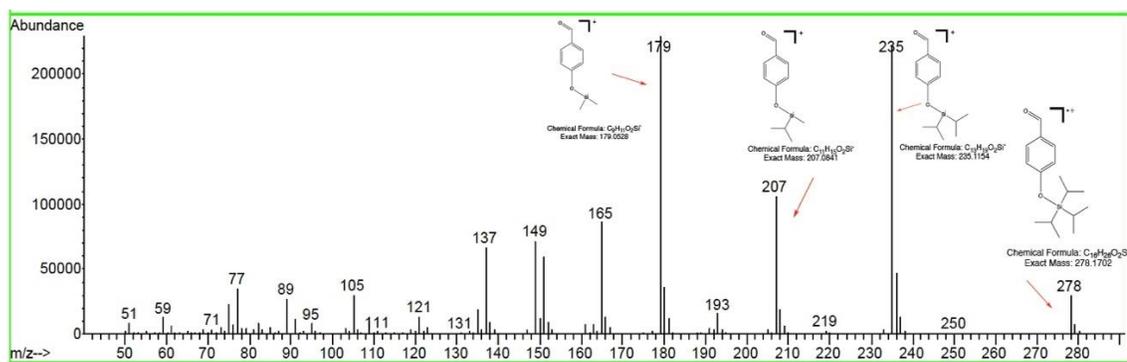
Figure 2.9 EI mass spectrum of H unit phenol-terminus precursor.

2.3.1.1   Charecterization of an Unsual Byproduct

The reaction for the synthesis of the H and G units had a high yield (~92%). However, the yield decreased considerably when attempting to synthesize the S unit, with the primary product being identified as an impurity or unreacted aldehyde. Monitoring of the reaction using thin-layer chromatography (TLC) proposed that the impurity was likely the unreacted aldehyde, as it exhibited the same retention factor (Rf) as the original aldehyde used in the reaction. Despite this, gas chromatography-mass spectrometry (GC-MS) analysis failed to detect the presence of the original aldehyde in either the reaction mixture or the isolated impurity. The intriguing finding obtained from GC-MS prompted us to conduct further investigations into this phenomenon. Subsequently, alternative analytical methods were employed to determine the identity of the enigmatic byproduct. The isolated byproduct was subjected to direct infusion analysis on a Q-Exactive orbitrap (ESI) positive mode mass spectrometer, which revealed the incorporation of an additional TIPS group into an S unit phenol-terminus precursor. The S unit precursor is a syringe aldehyde protected with a TIPS group with molecular mass of 338 amu. The calculated mass of $[M+H]^+$ is found to be m/z 339.1992. In the full scan mass spectrum of the byproduct a signal observed for S unit precursor at m/z 339.1986 (error: -1.77 ppm).Also, a base peak was observed at m/z 495.3321 which indicated addition of a TIPS group to the S unit phenol-terminus precursor (calc. mass m/z 495.3326, error: -1.01 ppm). It is noteworthy that the GC-MS analysis of the isolated byproduct produced results that were similar in terms of retention time and mass spectrum to those obtained for the S unit phenol-terminus precursor. However, the results obtained from TLC and Q-Exactive mass spectrometry suggested a different molecular composition for the byproduct. This led to

the hypothesis that the byproduct underwent a conversion to S unit precursor at the GC-MS inlet, as evidenced by the similar retention times observed for both compounds (Figure 2.11). The byproduct was also analyzed using HNMR and FTIR spectroscopy.
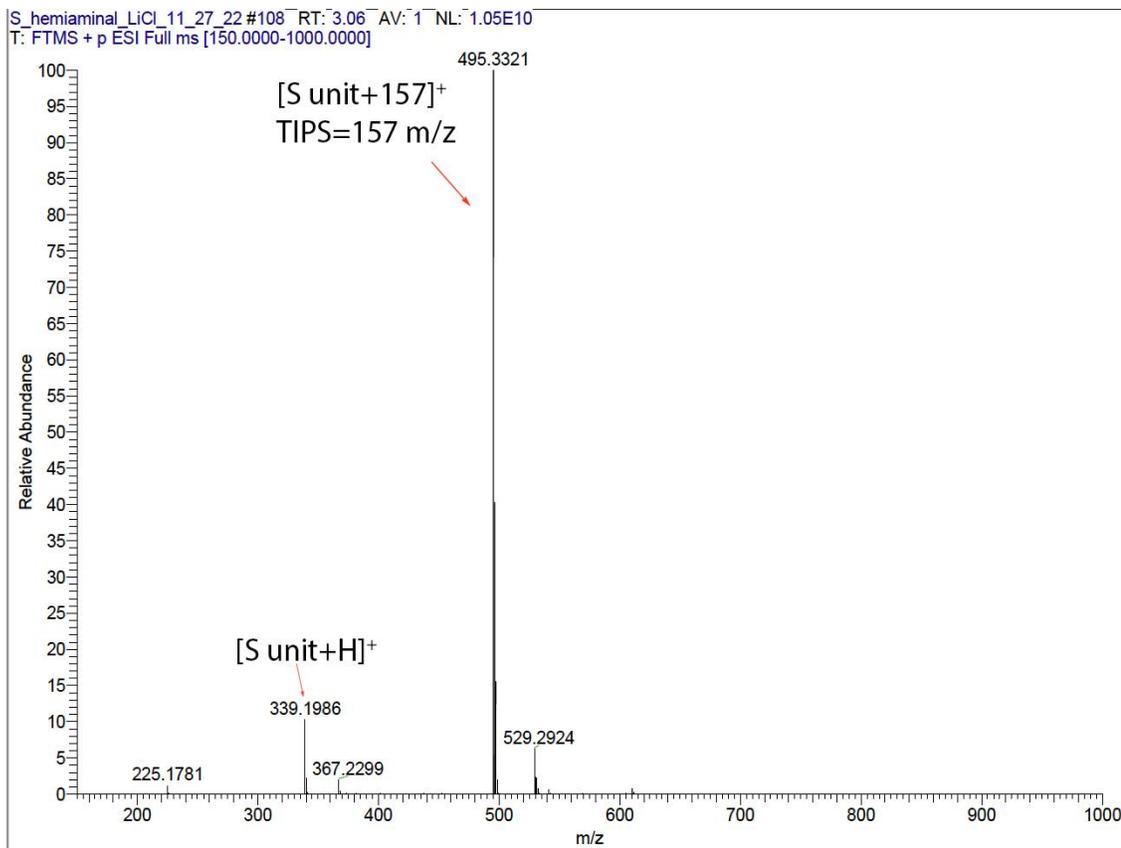
Figure 2.10 HRMS (ESI) Q-Exactive full scan mass spectrum for the byproduct. Calc. mass for $[C_{18}H_{31}O_4Si]^+$ ([M+H]) m/z 339.1992, obs. mass m/z 339.1986 (error=-1.77 ppm). Calc mass for $[C_{27}H_{51}O_4Si_2]^+$ ([M+157]) m/z 495.3326, obs. mass m/z 495.3321 (error=-1.01 ppm).

44
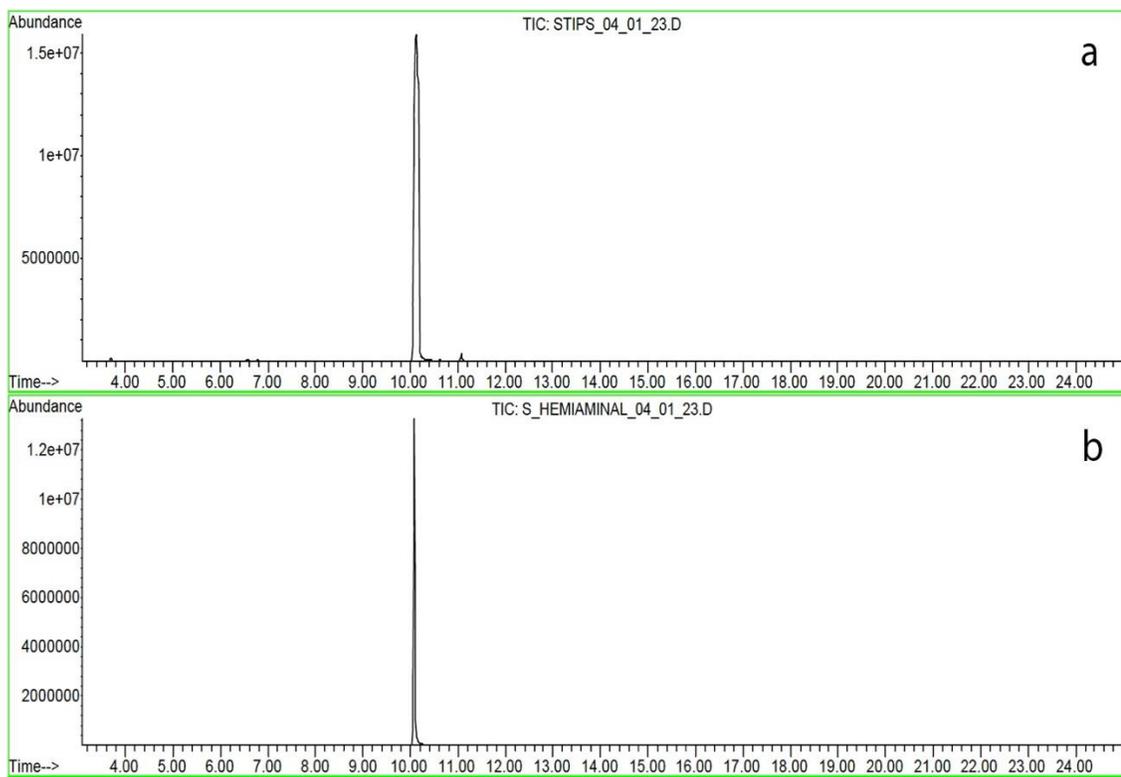
Figure 2.11 GC-MS chromatogram of S unit phenol-terminus precursor(a), and the byproduct (b).
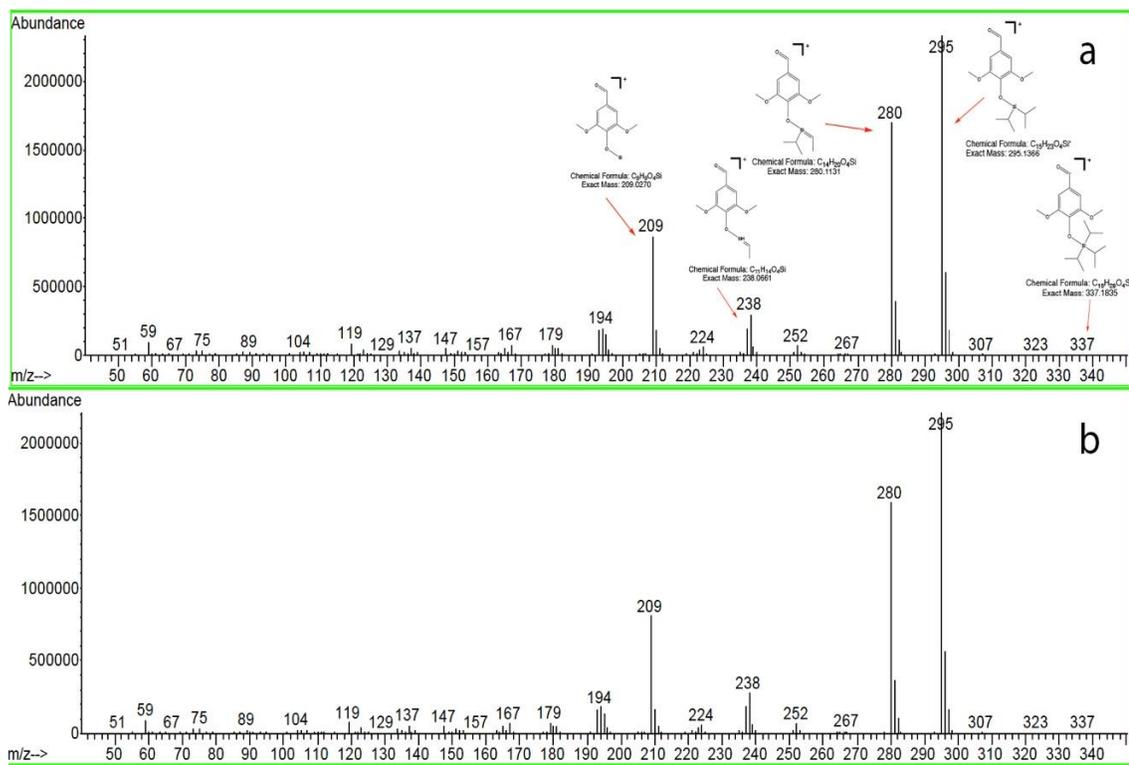
Figure 2.12 EI mass spectra of S unit phenol-terminus precursor(a), and the byproduct (b).

2.3.1.2   Nuclear Magnetic Resonance (NMR) Analysis

The isolated byproduct was dissolved in chloroform-d and was subject to 400 MHZ Bruker HNMR. The spectrum was calibrated base on the solvent peak at 7.24 ppm.

The NMR spectra of the synthesized compound showed characteristic signals for different protons. The hydrogens on the aromatic ring (1&2) exhibited overlap and appeared as a single peak with integration of 2 at approximately 6.5 ppm. The methoxy groups (3) exhibited a singlet at around 3.7 ppm, consistent with their expected chemical shifts. Two hydrogens on positions 7 and 8 appeared in the vinyl region at around 7 ppm. The hydrogen on position 9 experienced upfield shift due to its bonding with two nitrogen atoms, leading to deshielding, and appeared at around 7.5 ppm. The positions at 7,8, and 9 experienced long-range splitting in conjugated imidazole ring and formed doublet of doublet(dd). However, due the presence of electronegative nitrogen, the coupling constant is relatively low and appeared to be 1 Hz. Since the coupling constants for both splitting were equal, the dd appeared as triplets. The TIPS groups' hydrogens appeared at the beginning of the aliphatic region, showing splitting of quartets and doublets. The methyl groups on TIPS appeared as doublets and overlapped at 1 ppm, with an overall integration of 36 H. The quartet peaks of CH on the TIPS groups appeared at around 1-1.2 ppm with an overall integration of 6 H.

The NMR data integration confirmed the presence of two TIPS groups in the synthesized compound. The chemical shifts for the TIPS group and similar positions were correlated with a standard sample of S-TIPS, supporting the proposed structure.

Figure 2.13 HNMR spectrum of the byproduct

Figure 2.14 Characterized structure of by product. Positions numbers for HNMR analysis
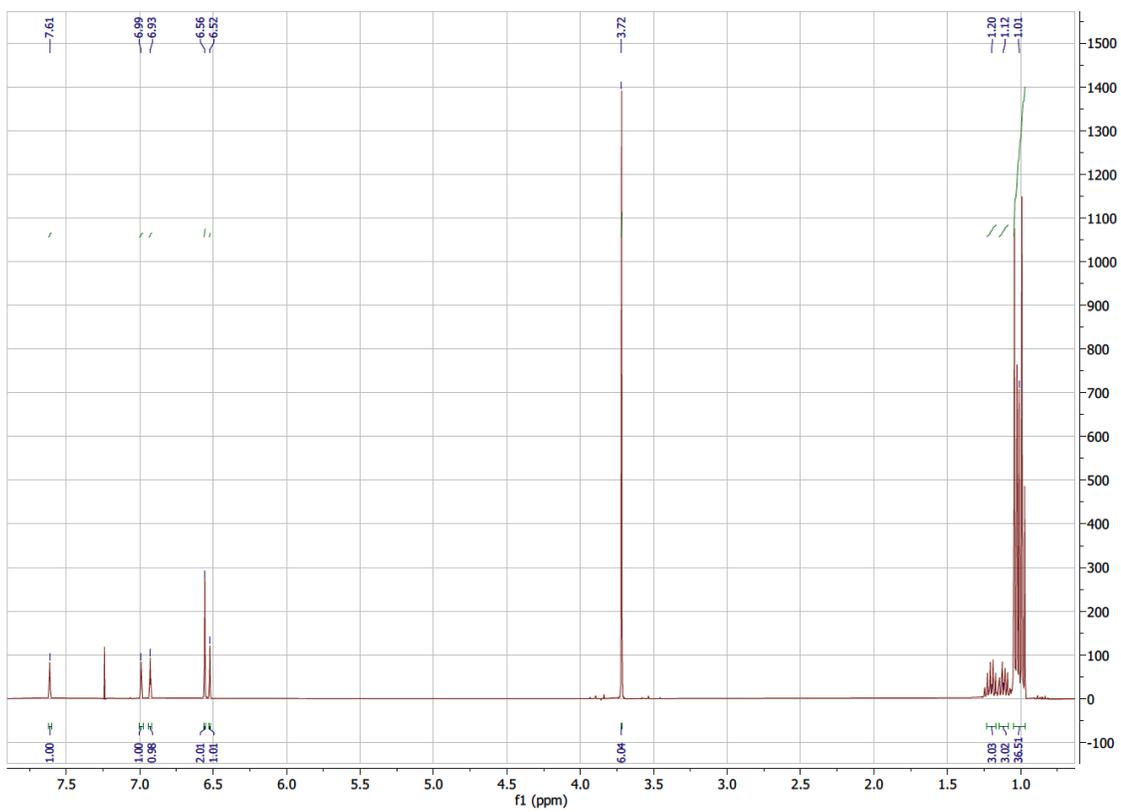
Table 2.1 HNMR results for the byproduct

position	δ_H, *(M; J in Hz)*
1	6.56 (s)
2	6.56 (s)
3	3.72 (s)
4	1.12-1.20 (q; 8.1)
5	1.01(d; 8.1)
6	6.52 (s)
7	6.93 (dd like t; 1)
8	6.99 (dd like t; 1)
9	7.61 (dd like t; 1)

2.3.1.3 Fourier Transform InfraRed (FTIR)

The proposed structure of the HNMR analysis was hypothesized to be the result of a nucleophilic attack of imidazole on the carbonyl group. The nucleophilic attack caused the carbonyl bond to open, and due to the presence of excess TIPSCl in the reaction vessel, the oxygen atom was stabilized by attaching to a TIPS group, ultimately leading to the proposed structure. To validate the proposed structure, FTIR was utilized.

The S unit precursor contained a carbonyl group, and according to the proposed mechanism, the byproduct should not contain this functional group. FTIR was employed to investigate the presence of carbonyl group in the byproduct.

The presence of carbonyl functional group in organic molecules can be detected by a characteristic peak in the FTIR spectrum, which is typically observed in the range of 1710-1685 cm^{-1} for alpha and beta unsaturated aldehydes[76, 77]. In the FTIR spectrum of the S unit precursor, a peak was observed at this region, indicating the presence of a carbonyl group. However, in the FTIR spectrum of the byproduct, this peak was absent, providing further evidence to support the proposed mechanism of the nucleophilic attack of imidazole. The FTIR spectroscopic analysis provided additional evidence to support the proposed structure of the byproduct.
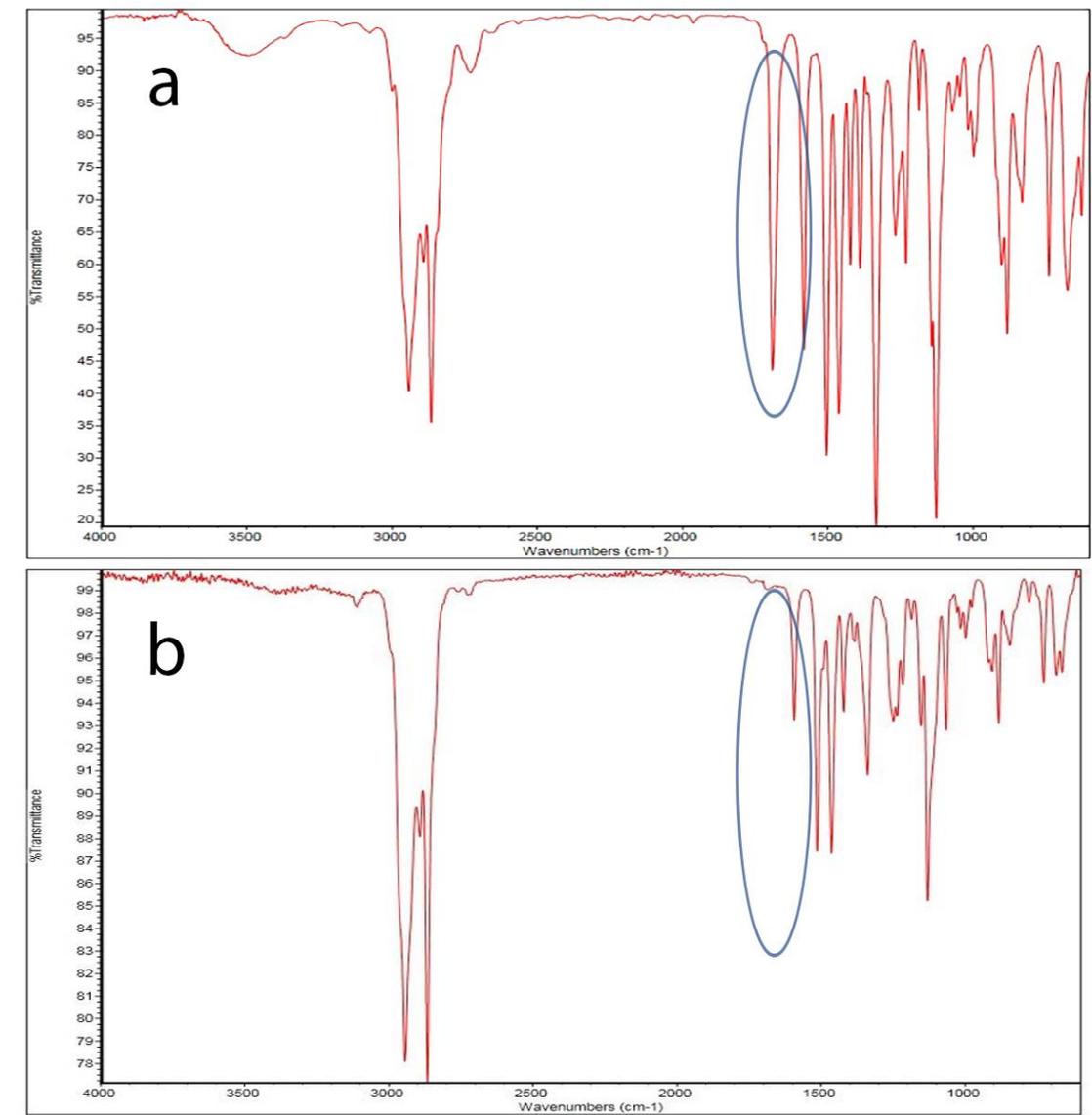
Figure 2.15 Comparison of FTIR spectra of S unit precursor (a) and the byproduct (b)

2.3.1.4   X-Ray Diffraction Crystallography (XRD)

X-ray diffraction (XRD) is a highly effective analytical technique that allows the determination of the crystal structure of a compound. This technique has the ability to provide valuable information regarding the arrangement of atoms in a crystal lattice, which enables the determination of the absolute configuration of the compound. In the present investigation, XRD was utilized to determine the crystal structure of the byproduct. To achieve this, different conditions of solvents and temperatures were applied to crystalize the byproduct. The obtained results of the XRD analysis confirmed the proposed structure of the byproduct.

The byproduct identified in this study was found to be a hemiaminal, resulting from a nucleophilic attack of imidazole to the carbonyl carbon of the aldehyde. Imidazole, a commonly used base in organic synthesis, was observed to be the source of this side reaction[78]. The reporting of this phenomenon can assist other researchers in taking this reaction into consideration during their experimental design.

In an effort to address this issue, alternative bases with lower nucleophilicity were investigated, including diisopropylethylamine (DIPEA), triethylamine (TEA), 4-dimethylamino pyridine (DMAP), and potassium carbonate. Although DIPEA, TEA, and DMAP produced the desired product without forming the hemiaminal byproduct, TEA was ultimately chosen due to its high yield and cost-effectiveness. These findings provide valuable insights for future synthetic endeavors and can help guide the selection of appropriate bases for similar reactions.
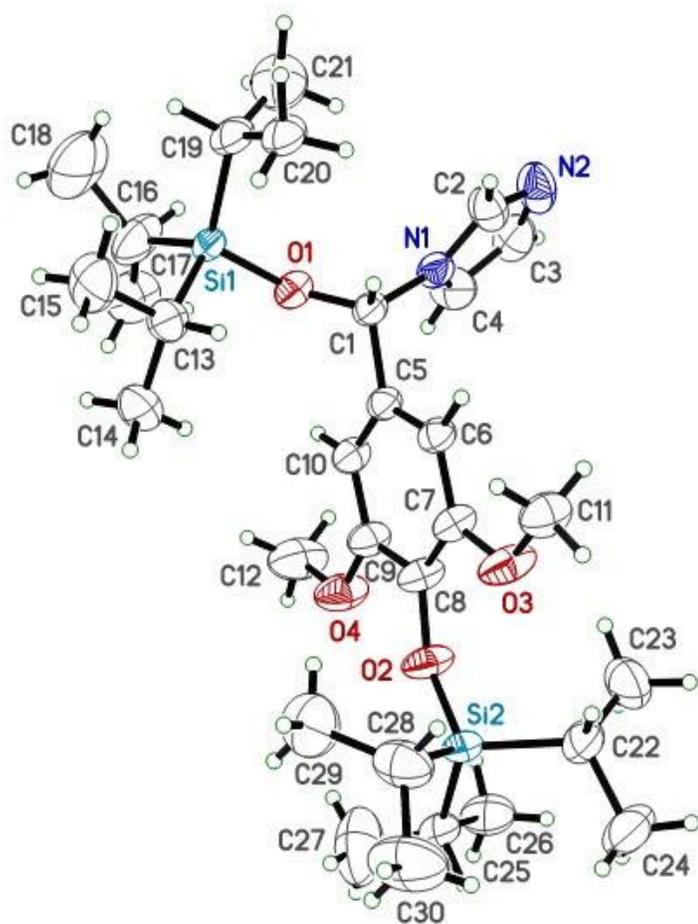
Figure 2.16 X-ray crystal structure of the byproduct.

2.3.2   Synthesis of Aliphatic-terminus Precursor

The aliphatic-terminus precursor serves as the final unit of the oligomer sequence and possesses an aliphatic chain that is similar to that found in natural lignin. In order to prepare this precursor for an aldol-like reaction, it is necessary to have an alpha carbon adjacent to the carbonyl functional group. To provide this alpha carbon, commercially available acid starting compound were first esterified. Specifically, p-coumaric acid, ferulic acid and sinapic acid were employed as starting compounds for the H, G and S units, respectively.

To achieve this, the acids were initially esterified with ethanol, and subsequently reacted with ethyl bromoacetate to provide a coupling site (alpha carbon) on the phenol end of the molecule. This approach allowed for the successful synthesis of the aliphatic-terminus precursor, which can be utilized in further chemical transformations for the production of lignin β-O-4 oligomers.

The synthesized compounds were characterized with GC-MS after derivatization with N, O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA). Figure 2.17 depicts the mass spectrum of ethyl p-coumarate, which represents the initial step in the synthesis of the aliphatic-terminus precursor. The molecular ion was identified as the base peak at m/z 264, and a characteristic alkyl loss was observed, as shown in Figure 2.17. A similar fragmentation pattern was observed for the G and S units, which exhibited a m/z 30 difference from the fragments observed for the H unit. Subsequently, the product obtained from this step is utilized in a reaction with ethyl bromoacetate to generate an alpha carbon suitable for aldol coupling.
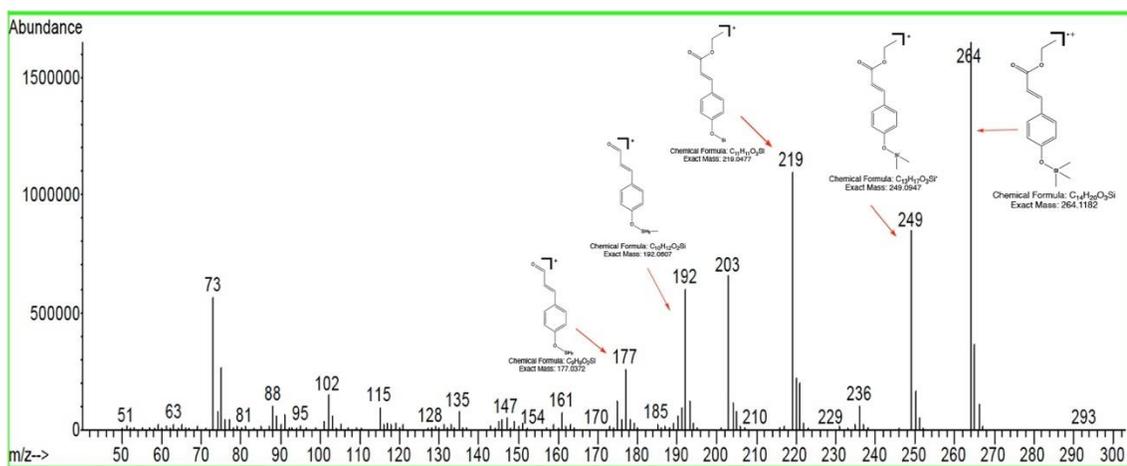
Figure 2.17 EI mass spectrum of ethyl p-coumarate.

The aliphatic-terminus precursor for the H unit exhibited a molecular ion as the base peak at m/z 278. The peak at m/z 250 corresponded to the loss of an ethyl group, which could arise from either of ethyl esters on both ends of the precursor, although the fragmentation of acid side of the precursor is depicted in Figure 2.18. Additionally, the precursor underwent fragmentation resulting in an ethoxy group loss, producing a fragment ion at m/z 233. Moreover, at m/z 206, the second ester group exhibited fragmentation and loss of an ethyl group. The precursor demonstrated multiple fragment ions, ultimately leading to the tropylium ion at m/z 91.
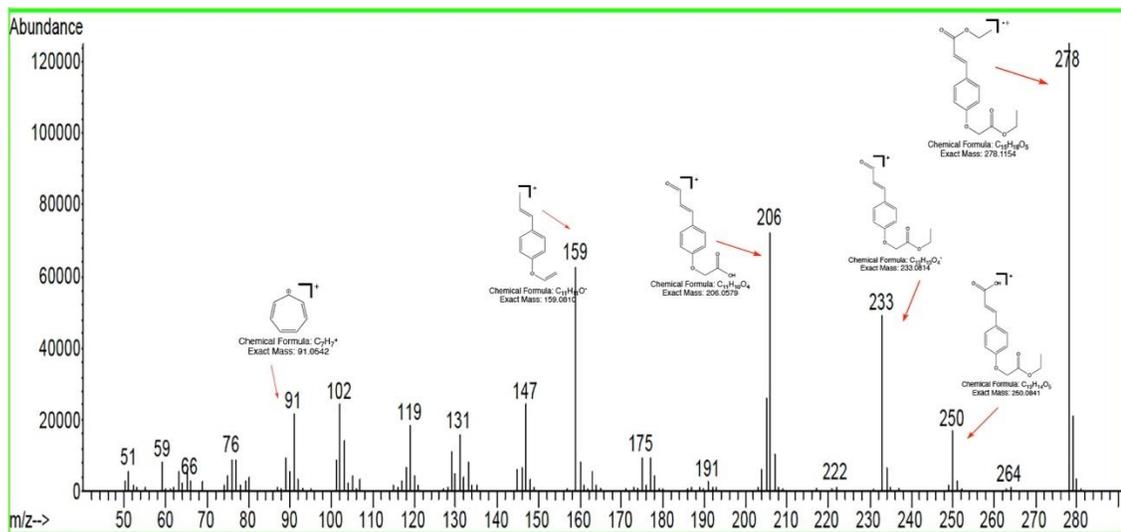
Figure 2.18 EI mass spectrum of H unit aliphatic-terminus.

2.3.3 Synthesis of Middle Ring Precursor

In order to synthesize sequences of β-O-4 trimers or hypothetical oligomers, a middle ring was prepared for each unit of H, G, and S. These precursors can be coupled with a phenol-terminus precursor and subsequently with an aliphatic-terminus precursor through an aldol reaction, followed by reduction and deprotection, to produce a β-O-4 model trimer. Hypothetically, the middle ring can undergo repeated coupling with additional middle rings, concluding in coupling with an aliphatic-terminus precursor to yield a β-O-4 oligomer. As mentioned in the previous section, an alpha carbon is necessary to carry out the aldol reaction. Since this is the middle ring, it also needs a carbonyl group for the subsequent aldol coupling reaction. Hence, in order to synthesize the precursor of the middle ring, a suitable aldehyde was utilized, namely 4-hydroxybenzaldehyde, vanillin, and syringe aldehyde for H, G, and S units, respectively. The aldehydes were reacted with ethyl bromoacetate to equip the precursor with an alpha carbon site on the phenol end of the ring. To avoid any side reactions on the aldehyde functional group of the ring, ethylene glycol was used as a protecting group. After each coupling reaction, the protecting group can be removed to expose a carbonyl group for the subsequent aldol addition reaction.
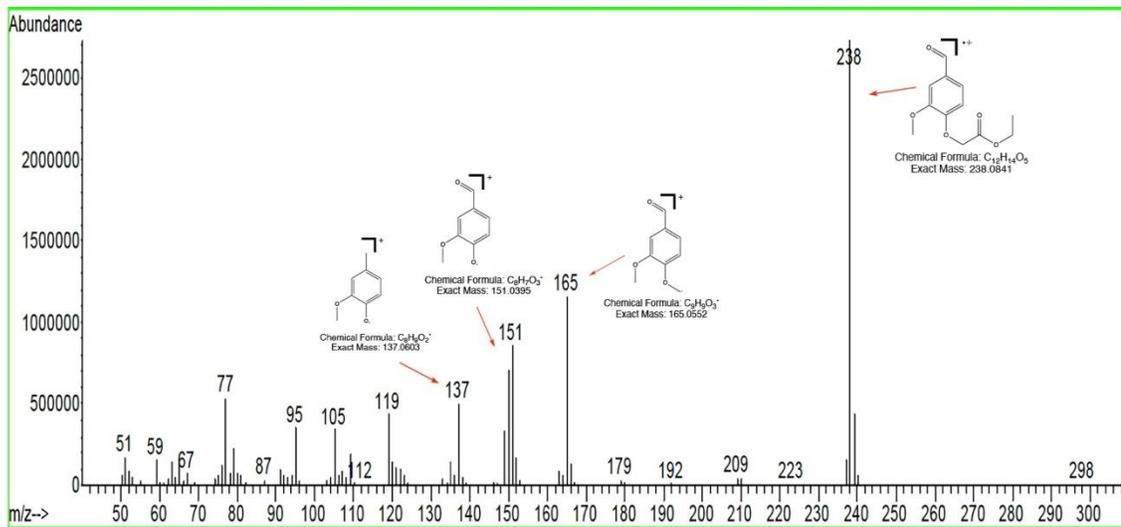
Figure 2.19 EI mass spectrum of G unit esterified vanillin for middle ring precursor.

In the synthesis of the middle ring precursor, the first step involved the esterification of an appropriate aldehyde. The G unit precursor synthesis started with esterification of vanillin. The EI mass spectrum of the esterified vanilin is shown in Figure 2.19. The EI ionization of all three units, H, G, and S provided the molecular ion, which for G unit appeared at m/z 238 (Figure 2.19). The most abundant fragment ion observed was at m/z 165, which resulted from the fragmentation of the bond between alpha carbon and the carbonyl group. The molecular ion also lost the ethyl acetate group, resulting in the m/z 151 fragment ion. This fragmentation pattern is indicative of the presence of the esterified vanillin in the G unit, which provides the alpha carbon site for the subsequent aldol addition reaction.

The second step to synthesize middle ring precursor was protecting the aldehyde functional group with ethylene glycol. The EI mass spectrum of G unit middle ring precursor is presented in Figure 2.20.

In the case of the middle ring precursor, the molecular ion was not observed. Instead, the mass of [M-H]$^+$ was detected at m/z 281 for G unit. The protection of the aldehyde with ethylene glycol resulted in the formation of a dioxolane group on the ring. It is noteworthy that in the literature, the [M-H]$^+$ fragment ion has been commonly reported for aldehydes, type three alcohols, and cyclic amines[79]. However, the EI fragmentation of 1,3-dioxolane has been reported in the literature with the base peak being [M-H]$^+$ at m/z 73 (as shown in Figure 2.21)[80, 81]. During the analysis of the middle ring precursor, the dioxolane ring was found to undergo hydrogen loss fragmentation. Notably, it was observed that the H unit predominantly exhibited [M-H]$^+$ fragment ion. Taking into account that the mass of molecular ion is equivalent to the isotopic mass of the discussed

fragment ion, the abundance of the molecular ion was relatively higher than the isotopic peak in G unit, indicating an increase in the abundance of the molecular ion. This trend led to the S unit exhibiting the molecular ion as more abundant than the $[M-H]^+$ fragment ion, and the $[M-H]^+$ still being at 60% height of the molecular ion. This was proposing that more electron donating methoxy groups stabilizing the molecular ion and increase its lifetime to be observed by the mass analyzer.
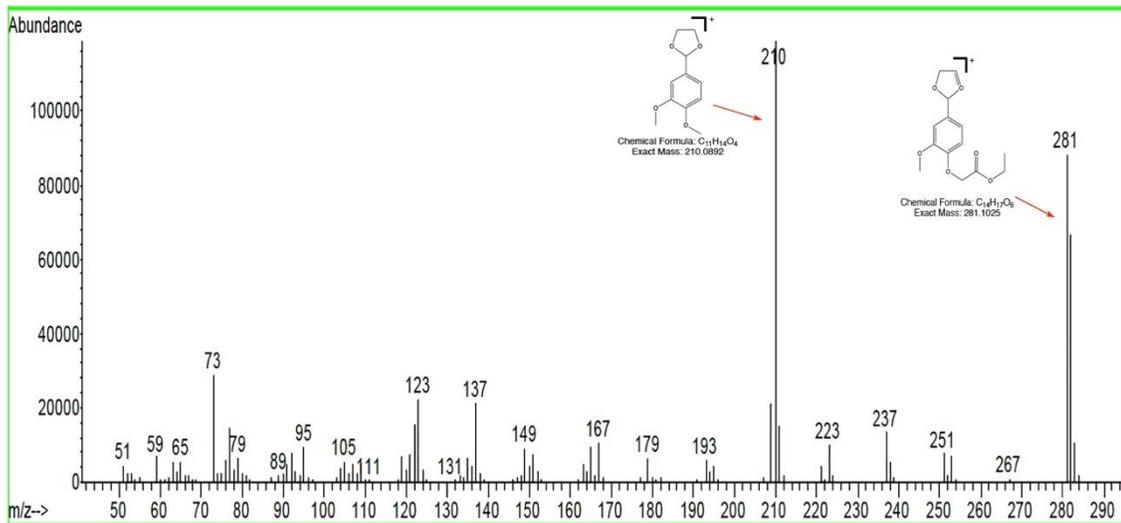
Figure 2.20 EI mass spectrum of G unit middle ring precursor.

1,3-Dioxolane
MASS SPECTRUM

NIST Chemistry WebBook (https://webbook.nist.gov/chemistry)
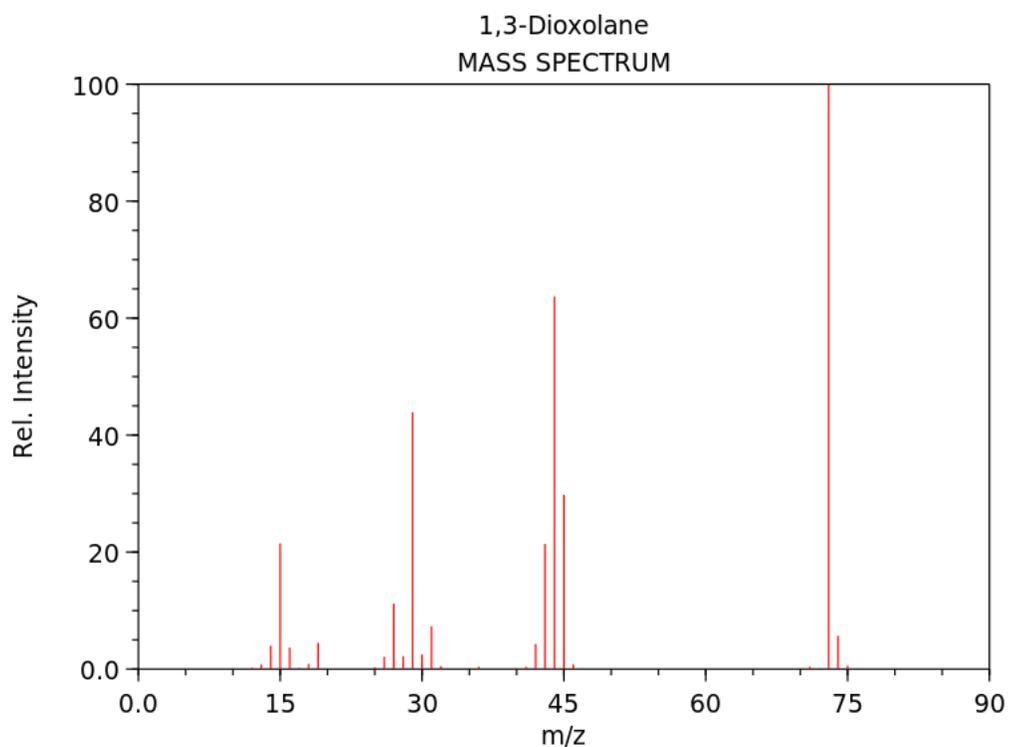
Figure 2.21 EI mass spectrum of 1,3-dioxolane[81].

64

2.3.4 Synthesis of Benzyl and Phenyl Modified G-βO4-G

The modification on G-βO4-G dimer was requested by our collaborator in department of chemical engineering to serve the purpose of a collaborative study[70, 82]. The modification was conducted on phenol side of the G-βO4-G dimer with benzyl and phenyl group.

The synthesis of the G-βO4-G dimer was achieved through $FeCl_3$-catalyzed oxidative coupling of G monolignols. For benzyl modification, the reaction was carried out by introducing benzyl bromide to G-βO4-G in the presence of potassium carbonate to yield benzyl-modified G-βO4-G dimer. The addition of phenyl group was conducted following Chan-Lam coupling of phenylboronic acid with G-βO4-G dimer. The synthesized dimers were characterized using GC-MS and sequence specific fragmentation investigation was carried out.

Through analysis, it was observed that the primary fragmentation pathway for both of the synthesized β-O-4 models was the cleavage of the bond between the alpha and beta carbon, leading to sequence-specific fragmentation. Figure 2.22 presents the EI fragmentation pattern of the benzyl-modified G-βO4-G dimer. The molecular ions of β-O-4 models are generally unstable; however, if the mass of the model is known, the molecular ion can be identified with low intensity. The benzyl-modified dimer yielded a low-abundance molecular ion of TMS-derivatized dimer at m/z 682. Upon fragmentation of α and β carbons, a sequence-specific fragment ion was observed for ring A at m/z 315. However, the fragment ion for ring B, resulting from this cleavage, was not detected, suggesting that ring A gained the charge while ring B fragmented as a neutral radical and remained undetected by the mass analyzer. Nonetheless, a ring B-specific fragment ion

appeared at m/z 324, which was common to both dimers and was indicative of ring B-specific fragmentation.

Figure 2.23 illustrates the EI mass spectrum of the phenyl-modified G-βO4-G dimer. Similar to the benzyl-modified dimer, the molecular ion of this model dimer was detected with low intensity. Moreover, the ring A specific fragments that were detected in the benzyl-modified dimer were also observed in the phenyl-modified dimer, albeit with a m/z 14 reduction compared to the values obtained for the benzyl-modified dimer, which resulted from structural differences of the two models. The fragment ions at m/z 301 and m/z 327 were ring A-specific fragment ions and equivalent to m/z 315 and m/z 341 ions observed in benzyl-modified dimer. The ring B-specific fragment ion was found to be identical for both the benzyl-modified and phenyl-modified dimers at m/z 324, as they share the same ring B structure.

In addition to the investigation of the modified G-βO4-G dimers, the fragmentation patterns of unmodified G-βO4-G dimers were also examined. It was found that the fragmentation patterns were consistent with those observed for the modified dimers, providing further evidence for the reliability of the observed fragmentation pathways. These fragment ions can be utilized as diagnostic tools for the GC-MS characterization of β-O-4 dimers, allowing for the identification and analysis of unknown lignin breakdown products. By utilizing the specific fragment ions for ring A and ring B, the structure of unknown lignin breakdown products can be predicted, providing insight into the composition and structure of natural lignin.
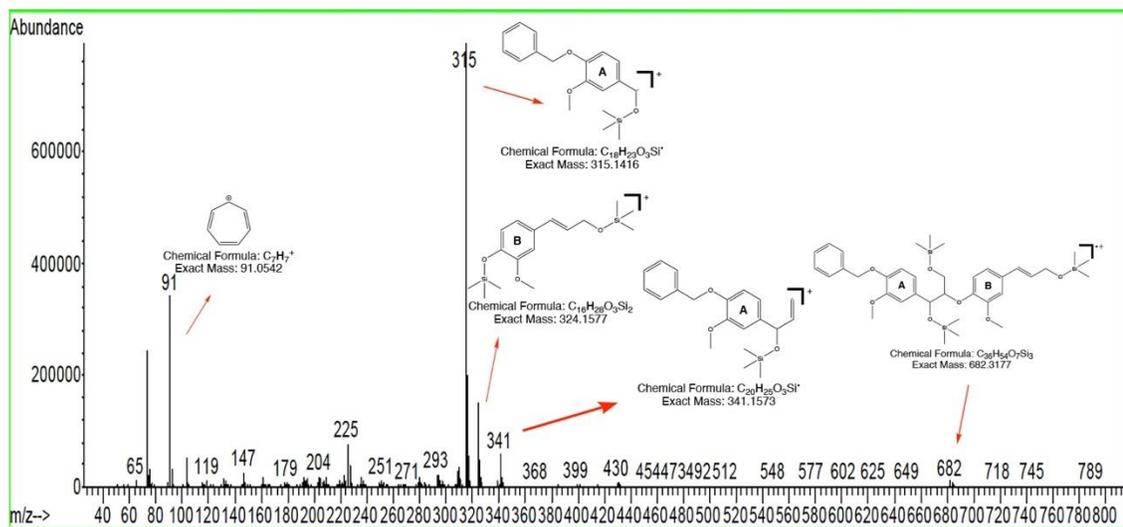
Figure 2.22 EI mass spectrum of benzyl-modified G-βO4-G dimer.
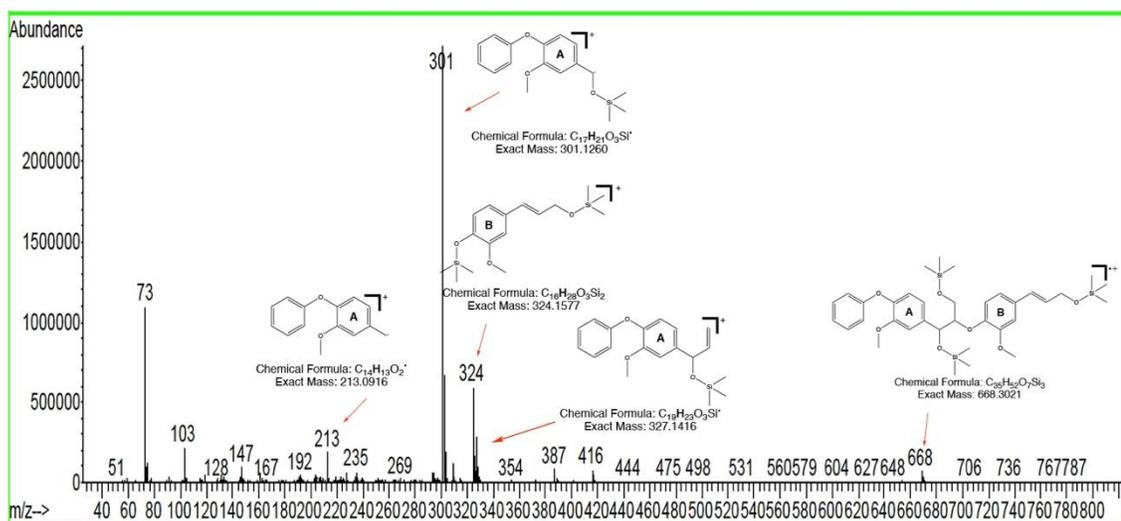
Figure 2.23 EI mass spectrum of phenyl-modified G-βO4-G dimer.
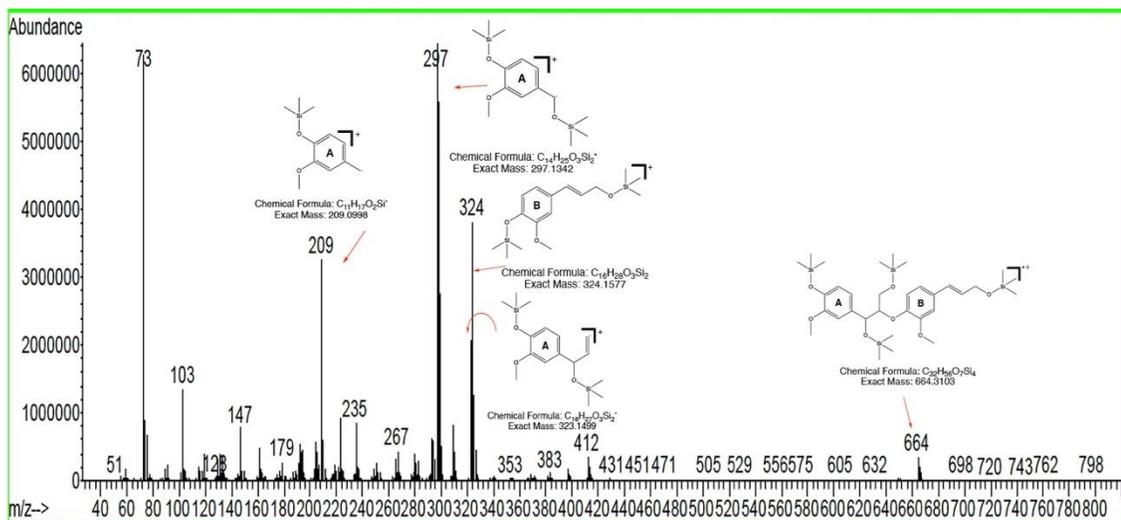
Figure 2.24 EI mass spectrum of G-βO4-G dimer.

2.4   Conclusion

In this study, we synthesized three discrete precursors for use in aldol reactions with the aim of producing β-O-4 model compounds. These precursors consist of a phenol-terminus precursor as the first monomeric unit, a middle ring precursor that can serve as a repeating unit for oligomer synthesis, and an aliphatic-terminus as the last monomeric unit. Each of these precursors were synthesized for H, G, and S units which vary in a methoxy group on the ring. These three types of precursors allow us to construct a variety of β-O-4 oligomers with a desired sequence using aldol reaction. The ability to synthesize sequence-specific β-O-4 oligomers in a controlled condition of aldol reaction represents a valuable tool to further enhance the structural analysis of lignin using mass spectrometry.

During the synthetic route of the S unit phenol-terminus precursor, an unexpected byproduct was observed. Gas chromatography-mass spectrometry analysis revealed that the byproduct had the same retention time and mass spectrum as the S unit phenol-terminus precursor. This suggests that the byproduct may have converted to the S unit precursor in the high temperature of the inlet. Further characterization of the byproduct was conducted using HRMS, HNMR, FTIR, and XRD, which showed that the byproduct was a hemiaminal formed by nucleophilic attack of an imidazole base on the carbonyl site, stabilized by excess TIPSCl in the reaction vessel. Imidazole is a commonly used base in organic synthesis. The discovery of this side reaction provides valuable insights for researchers working with imidazole or similar nucleophilic bases and can help them to design their experimental settings more effectively. As part of this investigation, alternative non-nucleophilic bases were evaluated, and triethylamine (TEA) was employed as a replacement for imidazole.

Furthermore, as per the request of our collaborator in the department of chemical engineering, we also synthesized structurally modified G-βO4-G dimers with both phenyl and benzyl groups. The synthesis process and EI fragmentation patterns of the modified G-βO4-G dimers were discussed in detail in this study and compared to fragmentation of unmodified G-βO4-G dimer. The results demonstrated that the G-βO4-G dimers produced structural-specific fragments in EI ionization. These fragmentation patterns can serve as a useful tool for GC-MS characterization of β-O-4 dimers.

The present study provides a detailed discussion of the Electron Ionization (EI) fragmentation patterns of various precursors. Notably, the $[M-H]^+$ fragment ion was observed for certain aldehydes and dioxolane group in the EI fragmentation of these precursors. The fragmentation patterns presented in this study can serve as a useful reference for the identification and characterization of similar model compounds, thereby advancing the structural analysis of lignin.

In this chapter, we have presented the synthesis and characterization of various lignin model compounds precursor to synthesize β-O-4 oligomers with desired sequences and designed to mimic the natural structure of lignin. Our study has also shed light on side reactions that can occur during the synthetic process, such as the formation of hemiaminals in the S unit phenol-terminus precursor. Furthermore, we have provided detailed EI fragmentation patterns for the synthesized model compound precursors and modified β-O-4 dimers that can be used for their characterization by GC-MS. Overall, the insights gained from this study will help advance our understanding of lignin and its mass spectral analysis, paving the way for the utilization and development mass spectrometry for the analysis of lignin.

CHAPTER 3. SYNTHESIS AND CHARACTERIZATION OF A MIXED LINKAGE LIGNIN TRIMER WITH β-O-4 AND β-5 BOND TYPES

3.1 Introduction

Lignin, a complex aromatic polymer, is an important component of plant cell walls and plays a crucial role in plant growth and development. As a result, lignin has become a subject of interest in various fields, including the biochemical, biomaterial, and pharmaceutical synthons. Although significant progress has been made in the development of analytical techniques for structural elucidation of lignin degradation products, further advancements are needed in this field.

Mass spectrometry has made significant progress, making it a powerful analytical tool for lignin analysis. High-resolution mass spectrometry (HRMS) in particular has undergone remarkable advancements that have significantly enhanced the accuracy, sensitivity, and comprehensiveness of characterization of lignin and its degradation products[83-87]. Despite significant advancements in high-resolution mass spectrometry (HRMS) technique, the ionization of lignin remains a challenge due to its complex and heterogeneous nature. Lignin is a large and polydisperse polymer that contains a variety of functional groups, such as phenolic hydroxyls, methoxys, and carboxylic acids, which can complicate ionization. Therefore, developing effective ionization methods for lignin analysis is critical in the field of mass spectrometry.

Negative ion mode is a common approach used in the analysis of lignin due to its ability to produce deprotonated molecular ions[88-90]. However, it should be noted that negative ion mode also has some limitations, such as low ionization efficiency, as well as a possible induced complexity in the mass spectrum[68]. Qi et al. investigated positive and negative ionization techniques and reported that they have observed higher number of

compounds detected in a mixture of lignin degradation products using negative ion mode ESI ionization compared to positive ion mode[91]. However, in their cationization, they have used ammonium formate to promote ionization and they were monitoring $[M+H]^+$, which does not provide intense signals in lignin compounds. The observation in our research group was that simple deprotonation results in charge-driven in-source fragmentation on the α-O-4 bond type, which causes a decrease in molecular ion intensity and an increase in complexity of the full scan spectrum[68]. Consequently, alternative ionization methods have been explored, and lithium cationization has been developed in previous studies. The findings from these studies revealed that β-O-4 and β-β dimers and oligomers can create stable cation adducts with lithium and generate lithiated fragment ions that provide sequence-specific information using tandem mass spectrometry analysis[52, 62, 71, 92, 93]. In the current work, we expanded our examination of lithium cationization by analyzing an oligomer with β-O-4 and β-5 bond types to achieve a structural informative fragmentation that can be used for analysis of lignin breakdown products.

An oligomer model trimer with β-O-4 and β-5 bond type has been synthesized. The synthesis was designed to retain the aliphatic chain with the double bond which has a critical role in chemical behavior of lignin. The oligomer was subject to fragmentation using Q-Exactive orbitrap mass spectrometer by high-energy collision dissociation (HCD) experiment. The trimer and fragments formed stable lithium adduct cation and a fragmentation pathway is developed. The fragment ion characterization can be applied on other lignin oligomers to elucidate the structure of unknown lignin degradation products.

3.2   Material and Methods

3.2.1   Chemicals

All chemicals were used without further purification. Vanillin, ferulic acid, ethanol, ammonium chloride, ferric chloride, ethyl bromoacetate, potassium carbonate, sodium bis(trimethyl silyl) amide (NaHMDS), and lithium chloride were purchased from Alfa Aesar (Ward Hill, MA, USA). Acetyl bromide, sodium sulfate, acetone, ethyl acetate, acetic anhydride, pyridine, optima acetonitrile, optima methanol, and optima water were purchased from Fisher Scientific (Fair Lawn, NJ, USA). Diisobutyl aluminum hydride (DIBAL-H, 1 M in methylene chloride) was purchased from Sigma-Aldrich (Saint Louis, MO, USA) and methylene chloride was purchased from VWR chemicals (Radnor, PA, USA).

3.2.2   Methods

3.2.2.1   Synthesis of Acylated Vanillin (Compound 2)

Vanillin (1g, 6.57 mmol) was dissolved in methylene chloride (20 mL) with a 1:20 w/v ratio. Acetic anhydride (3.11 mL, 32.9 mmol) and pyridine (1.06 mL, 13.1 mmol) were added to the solution with 1:5 and 1:2 molar ratios, respectively. The solution was stirred at 50°C for 3 hours. The product was purified with a silica gel gravity column (98% yield).

GC-MS $[M]^{\bullet+}$ calc. mass for $[C_{10}H_{10}O_4]^{\bullet+}$ m/z 194, obs. mass m/z 194.
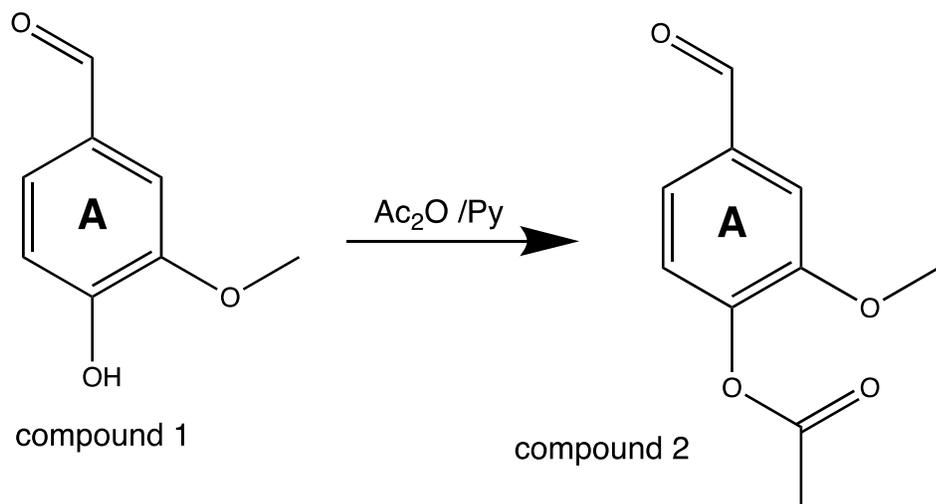
Figure 3.1 Synthesis of acylated vanillin.
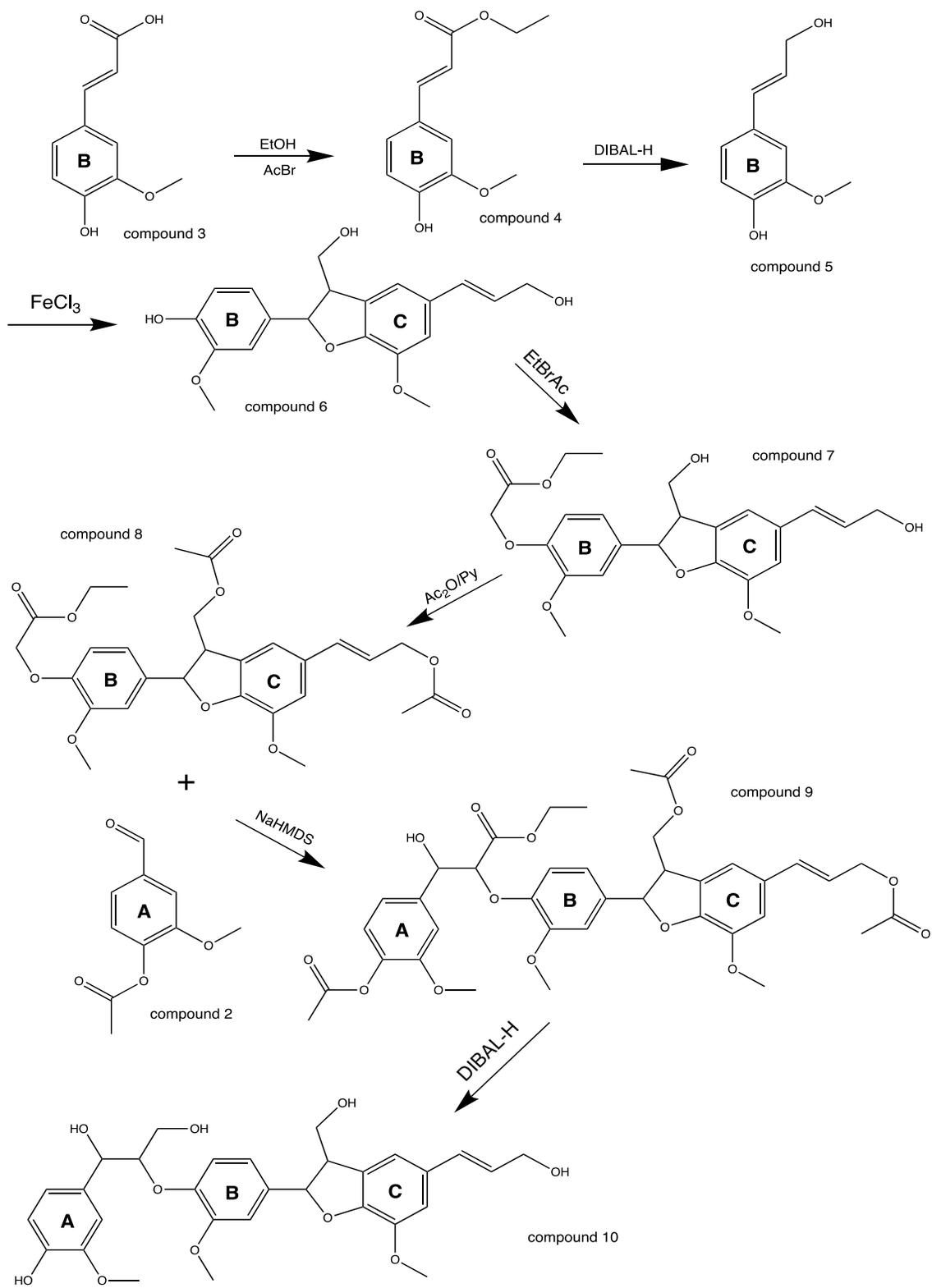
Figure 3.2 Synthetic route for G-βO4-G-β5-G trimer (compound 10).

3.2.2.2   Synthesis of Compound 8

Compound 8 as the precursor for the aldol reaction was synthesized in 5 steps.

3.2.2.2.1   SYNTHESIS OF ETHYL FERULATE

Ethyl ferulate was synthesized as explained in Section 2.2.2.2.1(97% yield).

GC-MS $[M]^{\bullet+}$ calc. mass for TMS derivatized compound $[C_{15}H_{22}O_4Si]^{\bullet+}$ m/z 294, obs. mass m/z 294.

3.2.2.2.2   SYNTHESIS OF G MONOLIGNOL (COMPOUND 5)

All glassware were kept at 120°C overnight. Ethyl ferulate (3 g, 13.5 mmol) was dissolved in methylene chloride (100 mL).The solution was then sealed under nitrogen and kept at -78°C. Diisobutyl aluminum hydride (DIBAL-H) 1M in methylene chloride (40.5 mL) was added dropwise to the solution with a syringe under nitrogen by 1:3 molar ratio. The solution was then stirred for three hours after DIBAL-H was added, quenched with saturated ammonium chloride (100 mL), and extracted with ethyl acetate (150 mL × 3 times) and purified with silica gel gravity column (95% yield)

GC-MS $[M]^{\bullet+}$ calc. mass for TMS derivatized compound $[C_{16}H_{28}O_3Si_2]^{\bullet+}$ m/z 324, obs. mass m/z 324.

3.2.2.2.3   SYNTHESIS OF G-β5-G DIMER (COMPOUND 6)

The β-5 dimer was synthesized with ferric chloride oxidative coupling reported by Lancefield et al. with some modification[94]. Briefly, G monolignol (500 mg, 2.78 mmol)

was dissolved in acetone (10 mL), and then water (50 mL) was added to the solution. Ferric chloride (427.5 mg, 2.64 mmol) with a 1:0.95 molar ratio was dissolved in water (30 mL) and added dropwise to the G monolignol solution. The reaction mixture quenched one hour after the addition of ferric chloride by extraction with ethyl acetate (100 mL × 3 times). G-β5-G dimer was isolated with a silica gel column from the product mixture (22% yield).

GC-MS $[M]^{\bullet+}$ calc. mass for TMS derivatized compound $[C_{29}H_{46}O_6Si_3]^{\bullet+}$ m/z 574, obs. mass m/z 574.

3.2.2.2.4 SYNTHESIS OF COMPOUND 7

The addition of ethyl bromoacetate to the phenol end of G-β5-G dimer was conducted as reported in Section 2.2.2.2.2. G-β5-G dimer (800 mg, 2.23 mmol) dissolved in acetone (50 mL). Then, ethyl bromoacetate (372 μL, 3.35 mmol) and potassium carbonate (463 mg, 3.35 mmol) were added to the solution and refluxed for five hours. The reaction was monitored by microextraction and injection in GC-MS, as the Rf of compound 7 was close to the starting dimer. The reaction mixture decanted and transferred to a clean glassware to remove excess of potassium carbonate and concentrated to ~20 mL solution. The product was extracted with ethyl acetate (100 mL × 3 times). Compound 7 was purified with silica gel gravimetry column (88% yield).

GC-MS $[M]^{\bullet+}$ calc. mass for TMS derivatized compound $[C_{30}H_{44}O_8Si_2]^{\bullet+}$ m/z 588, obs. mass m/z 588.

3.2.2.2.5 SYNTHESIS OF COMPOUND 8

The aliphatic hydroxyl group on the compound 7 was protected with acylation by acetic anhydride and potassium carbonate. Compound 7 (800 mg, 1.80 mmol) dissolved in acetone (50 mL). Acetic anhydride (852 μL, 9.01 mmol) and potassium carbonate (747 μL, 5.41 mmol) was added to the solution. The solution was stirred with a magnet stirrer at 50°C for an hour. The reaction was monitored with TLC plate and was decanted to a clean glassware to remove excess potassium carbonate and concentrated to ~20mL. The product was extracted with ethyl acetate (100 mL × 3 times) and saturated ammonium chloride (70 mL) and purified with silica gel gravity column (83% yield).

LTQ (ESI- positive ion mode) $[M+Li]^+$ calc. for $[C_{28}H_{32}O_{10}Li]^+$ m/z 535.2156, obs. mass m/z 535.2.

3.2.2.3 Synthesis of Compound 9

All glassware were kept at 120°C overnight. Compound 2 (36.7 mg, 0.19 mmol) and compound 8 (50 mg, 0.095 mmol) were dissolved in anhydrous THF (2 mL for each compound) in separate dry pear-shape flask and purged with nitrogen. Sodium bis(trimethylsilyl)amide (NaHMDS) (86.7 mg, 0.43 mmol) base weighed with 1:5 molar ratio to compound 8 and dissolved in anhydrous THF (5 mL) in a 3-neck flask and purged with nitrogen. All reagents were sealed with rubber septum. Compound 8 solution was transferred to an addition funnel with a stainless-steel syringe needle (gauge 22, L 12in, Sigma-Aldrich, Saint Louis, MO USA) under nitrogen. The solution was added dropwise to the reaction vessel at -78°C under N2 by addition funnel with pressure equalizer to a 3-

neck flask that contains the base solution. The reaction mixture was stirred for one hour, then compound 2 solution was added dropwise with the addition funnel and was stirred for 4 hours until compound 8 was completely reacted. The reaction was quenched with saturated ammonium chloride (10 mL) and extracted with ethyl acetate (10 mL × 3 times). The product used for next step without purification.

LTQ (ESI- positive ion mode) [M+Li]$^+$ calc. for [C$_{38}$H$_{42}$O$_{14}$Li]$^+$ m/z 729.2735, obs. mass m/z 729.2.

3.2.2.4 Synthesis of Compound 10

The product from the last step was reduced with DIBAL-H to yield the final trimer (compound 10). The reaction mixture from previous step was used without any purification. DIBAL-H (0.28 mL, 0.28 mmol) was measured by assuming the conversion of all the B ring into the trimer and used that mass to measure 1:3 molar ratio for DIBAL-H. The product was purified with silica gel gravity column and then the fraction is further purified with HPLC.

Q-Exactive (ESI-Orbitrap) [M+Li]$^+$ calc. for [C$_{30}$H$_{34}$O$_{10}$Li]$^+$ m/z 561.2312, obs. mass m/z 561.2264. Error (ppm)= -8.5526.

^{1}HNMR (600 MHz, CD$_3$COCD$_3$): δ 7.10 (d, J = 1.5 Hz, 1H), 6.76 (d, J = 7.3 Hz, 1H), 6.89 (d, J = 7.3 Hz, 1H), 4.88 (d, J = 5.5 Hz, 1H), 4.31 (q, J = 5.2 Hz, 1H), 3.69-3.80 (m, 2H), 3.81 (s, 3H), 7.05 (br s, 1H), 6.89 (d, J = 6.6 Hz, 1H), 6.95 (d, J = 6.6 Hz, 1H), 5.59 (d, J = 6.2 Hz, 1H), 3.51 (br q, J = 6.3 Hz, 1H), 3.80-3.88 (m, 2H), 3.87 (s, 3H), 6.95 (d, J = 1.5 Hz, 1H), 6.97 (s, 1H), 6.53 (d, J = 15.9 Hz, 1H), 6.25 (dt, J = 15.8, 5.5 Hz, 1H), 4.19 (d, J = 5.5 Hz, 2H), 3.87 (s, 3H), 7.44 (s, 1H), 4.56 (s, 1H), 2.85 (s, 3H)

^{13}CNMR (150 MHz, CD$_3$COCD$_3$): δ 134.78, 111.92, 152.15, 148.43, 115.57, 121.01, 74.25, 87.07, 62.29, 56.71, 137.73, 111.75, 152.29, 149.26, 119.68, 119.64, 88.75, 55.36, 65.12, 56.88, 132.56, 112.21, 145.70, 149.38, 130.73, 116.59, 130.96, 128.96, 63.81, 56.89 ppm.

3.2.3 Instruments

3.2.3.1 Nuclear Magnetic Resonance (NMR)

NMR spectra were performed using a Bruker (Billerica, MA, USA) Ascend 600 (^{1}H, 600 MHz; ^{13}C, 150 MHz) spectrometer. Purified compound 10 was dissolved in 500 μL acetone-d^6 and the HNMR and CNMR spectra were calibrated based on the solvent peak at 2.05 and 206.7 ppm, respectively.

3.2.3.2 High Performance Liquid Chromatography (HPLC)

Semipreparative HPLC was used on an Agilent (Santa Clara, CA, USA) 1260 Infinity II HPLC equipped with a diode array detector (DAD) using a Phenomenex (Torrance, CA, USA) Gemini C18 column (250 × 10 mm, 5 μm). HPLC solvents used were of ACS grade and purchased from the VWR (Radnor, PA, USA).

Mixture of compounds (15.2 mg) was dissolved in methanol and subjected to semipreparative HPLC (solvent A: H$_2$O/0.025% TFA; solvent B: MeOH; flow rate: 5.0 mL min-1; 0−2 min, 20% B; 2−4 min, 20-40% B; 4−8 min, 40-55% B; 8−18 min, 55% B; 18-25 min, 55-100% B; 25−27 min, 100% B; 27−28 min, 100-20% B; 28−30 min, 20%

B.) afforded Fraction A (compound 10, 4.3 mg, Rt: 11.45 min) and two other fractions (Fraction B: 2.5 mg at 12.10 min and Fraction C: 3.6 mg 13.05 min).

3.2.3.3   Gas Chromatography-Mass Spectrometry

The GC-MS analysis was conducted on an Agilent (Santa Clara, CA, USA) 5973 MSD equipped with HP 6890 GC and HP 7683 injector controlled by ChemStation D.03.00611. A DB-5HT. A GC column with 15m length, internal diameter of 250 μm and film thickness of 0.1μm was used. The injector temperature was set to 250 °C. The oven temperature was set at 100 °C and held for 3 minutes, then ramped at 15 °C/min to 280 °C and held for 10 minutes for a total method time of 25 minutes. An injection split ratio of 50:1 was used for analysis.

3.2.3.4   Linear Ion Trap Quadrupole Mass Spectrometer

Sample preparation for LTQ spectrometer was achieved by preparing 0.5 mg/mL solution of the analyte in 1:1 mixture of optima grade acetonitrile and aqueous LiCl (10 mM). Analytes introduce to mass spectrometer with direct infusion using a syringe pump with 3 μl/min flow rate.

The Thermo Finnigan LTQ mass spectrometer (San Jose, CA, USA) was equipped with an ESI source with a spray voltage of 3.80 kV, capillary temperature of 250°C and sheath gas flow rate of 2 (arbitrary unit) was applied.

3.2.3.5   Q-Exactive Orbitrap Mass Spectrometer

Sample preparation was the same as LTQ mass spectrometer with lower concentration of 0.2 mg/mL in 1:1 mixture of optima grade acetonitrile and aqueous LiCl (10 mM).

Accurate mass and higher energy collisional dissociation (HCD) experiments, conducted on a Thermo Scientific Q-Exactive Orbitrap (Waltham, MA, USA). The Q-Exactive was equipped with a HESI probe with a spray voltage of 3.8 kV. The inlet temperature was set at 225°C and a sheath auxiliary gas flow of 2 (arbitrary unit) was used for analysis. The HCD experiment on the analyte was conducted at normalized collision energy (NCE) of 30%-39%, and nitrogen as collision gas.

3.3   Discussion

The synthesis of model compounds has greatly contributed to the elucidation of the structure of lignin. The majority of model compounds that have been developed struggle to mimic the actual structure of lignin. There is no doubt that the double bond on the aliphatic chain of lignin plays an important role in the chemical behavior of lignin. The resonance between the double bond and the ring leads to the occurrence of different bond types in lignin polymer. This important fact, however, is largely absent from most of the model compounds that have been reported in the literature.

The most abundant bond type in lignin is β-O-4 bond. Therefore, most of the studies have been conducted on this type of bond. By using a more complex model compound with β-O-4 bonds as well as other types of bonds, we can more accurately analyze native

lignin. In this study, a trimer model compound containing β-O-4 and β-5 types has been synthesized.

To synthesize the trimer model compound, ferulic acid was reacted with acetyl bromide and ethanol to make ethyl ferulate and reduced with DIBAL-H to make coniferyl alcohol, also known as G monolignol. The monolignol was then subject to oxidative coupling with ferric chloride. Coupling of G monolignol results in β-O-4, β-5 and β-β dimers, where the G-β5-G dimer was isolated and purified for the synthesis of the trimer.

The β-O-4 bond was generated with an aldol reaction which require an alpha carbon and a carbonyl group. The phenol end precursor for this coupling was acylated vanillin while for the aliphatic end precursor, the G-β5-G dimer required two steps of reactions in order to be ready for the coupling. As part of the preparation of the alpha carbon on the precursor for the coupling, G-β5-G dimer was reacted with ethyl bromoacetate. This step provided the alpha carbon to the carbonyl, required for aldol coupling by bonding of ethyl acetate to the phenol end of the β-5 dimer. Protection of aliphatic hydroxyl groups on the dimer was also required to prevent interference with the coupling reaction. In order to achieve this, the product from the previous step was acylated in acetone with acetic anhydride and potassium carbonate to prepare the precursor for the aldol reaction. To form the β-O-4 bond, the aldol reaction was carried out under nitrogen and at -78°C. When the product of the coupling was retrieved, the acylation protecting groups were removed by reduction with DIBAL-H to obtain the final trimer. The trimer was purified using silica gel gravimetry chromatography and further purified by semipreparative HPLC. The pure trimer was subject to 600 MHz NMR to obtain HNMR, CNMR spectra, as well as 2D HMBC and HSQC spectra for structural analysis.

3.3.1 Characterization of G-βO4-G-β5-G Trimer by NMR

Due to the polar nature of the synthesized G-βO4-G-β5-G trimer, acetone-d^6 was selected as the NMR solvent. The trimer was subject to 600 MHz NMR spectrometer to obtain HNMR, CNMR and 2-D HSQC and HMBC. The result was compared to NMR directory provided by John Ralph and Sally Ralph for lignin derived compounds[95]. In the NMR directory, the dimers with β-O-4 and β-5 bonds were investigated and the moieties of the dimer which was repeated in the trimer structure were used as a reference to find the peaks.

Heteronuclear Single Quantum Coherence (HSQC) shows the single bond correlation of carbon and hydrogen. By using HSQC, it was possible to assign the observed proton peak in HNMR to the corresponding carbon peak in CNMR.

Heteronuclear Multiple Bond Correlation (HMBC) provides the correlation between hydrogen and carbon that are separated by two, three, or even four bonds in a conjugated system. Direct carbon-hydrogen bonds are suppressed and usually not observed in HMBC[96]. The correlation between carbon-hydrogen pairs assigned from HSQC is generally investigated in two ways: 1) by examining the correlations between each hydrogen peak and its neighboring carbons or 2) by examining the correlations between each carbon and its neighboring hydrogens. The correlation from carbon to their neighboring hydrogens was used to assign the peaks.

The list of the HNMR and CNMR peaks and their HMBC correlations are provided in Table 3.1
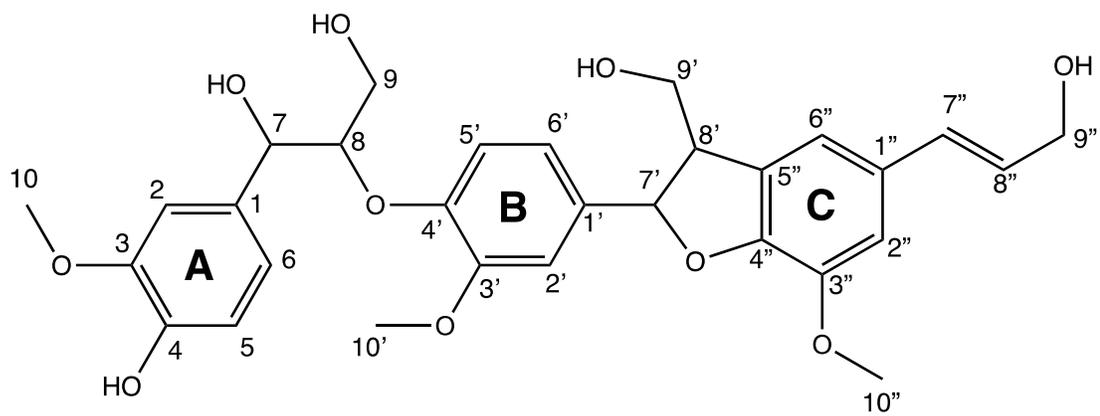
Figure 3.3 Chemical structure of G-βO4-G-β5-G trimer (compound 10). Carbons are numbered for NMR analysis.

Table 3.1 HNMR and CNMR results for G-βO4-G-β5-G trimer (compound 10) and HMBC correlations.

Position	δ_C, type	δ_H, (M; J in Hz)	HMBC
1	134.78, C		2, 5, 7, 8
2	111.92, CH	7.10 (d; 1.5)	7
3	152.15, C		10
4	148.43, C		5, 10
5	115.57, CH	6.76 (d; 7.3)	6
6	121.01, CH	6.89 (d; 7.3)	5, 2, 7
7	74.25, CH	4.88 (d; 5.5)	2, 5, 6, 8, 9
8	87.07, CH	4.31 (q; 5.2)	7, 9
9	62.29, CH_2	3.69-3.80 (m)	7, 8
10	56.71, CH_3	3.81 (s)	
1′	137.73, C		6′, 7′, 8′
2′	111.75, CH	7.05 (br s)	7′, 6′
3′	152.29, C		2′, 6′, 10′
4′	149.26, C		8,
5′	119.68, CH	6.89 (d; 6.6)	2′, 7′
6′	119.64, CH	6.95 (d; 6.6)	2′
7′	88.75, CH	5.59 (d; 6.2)	8′, 9′
8′	55.36, CH	3.51 (br q; 6.3)	7′, 9′, 6″
9′	65.12, CH_2	3.80-3.88 (m)	8′, 7′
10′	56.88, CH_3	3.87 (s)	
1″	132.56, C		7″, 8″, 9″
2″	112.21, CH	6.95 (d; 1.5)	6″, 7″
3″	145.70, C		2″, 10″
4″	149.38, C		7′, 8′, 2″, 6″
5″	130.73, C		8′, 7′, 9′, 6″
6″	116.59, CH	6.97 (s)	2″, 7″
7″	130.96, CH	6.53 (d;15.9)	2″, 6″, 9″
8″	128.96, CH	6.25 (dt; 15.8, 5.5)	9″
9″	63.81, CH_2	4.19 (d; 5.5)	7″, 8″
10″	56.89, CH_3	3.87 (s)	
OH-4		7.44	
OH-7		4.56	
9,9′,9″ OH		2.85	

The peaks for HNMR and CNMR spectra were calibrated using the solvent peak at 2.05 and 206.7 ppm, respectively. The numbering of the carbons in the structure was assigned using the conventional method that is well-known to lignin chemists.

As it can be observed on the structure of β-O-4 bond, there are 2 chiral centers on these types of bonding. As a result of these two chiral centers, four different isomers can be formed, which in turn can be divided into two groups of diastereomers known as erythro and threo. According to the lignin derived compounds NMR directory, the chemical shift for these two diastereomers have a slight difference. Considering that our purified synthesized trimer consists of these two diastereomers, most of the chemical shifts for carbon and hydrogen are split into two peaks that are approximately 0.02 ppm apart. This is coming from the difference between the chemical shifts of erythro and threo diastereomers.
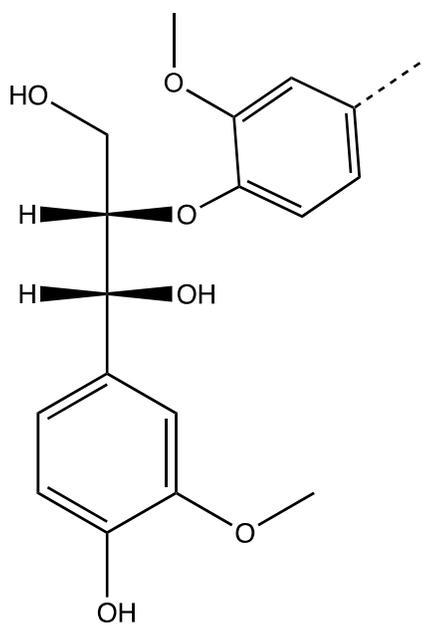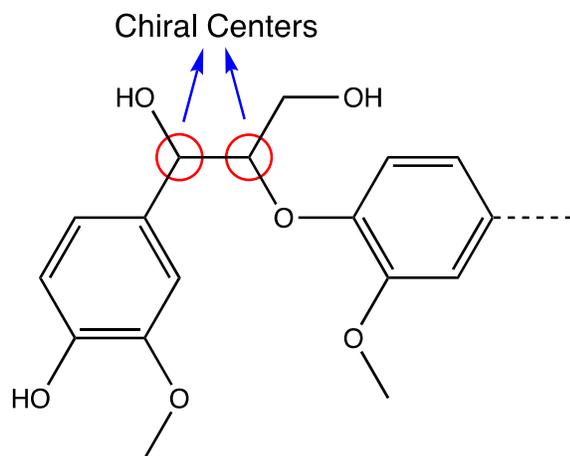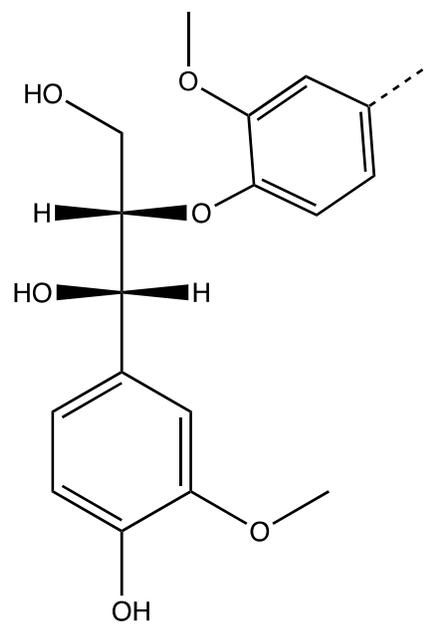
Figure 3.4 Diastereomer pairs in compounds with β-O-4 bond type.
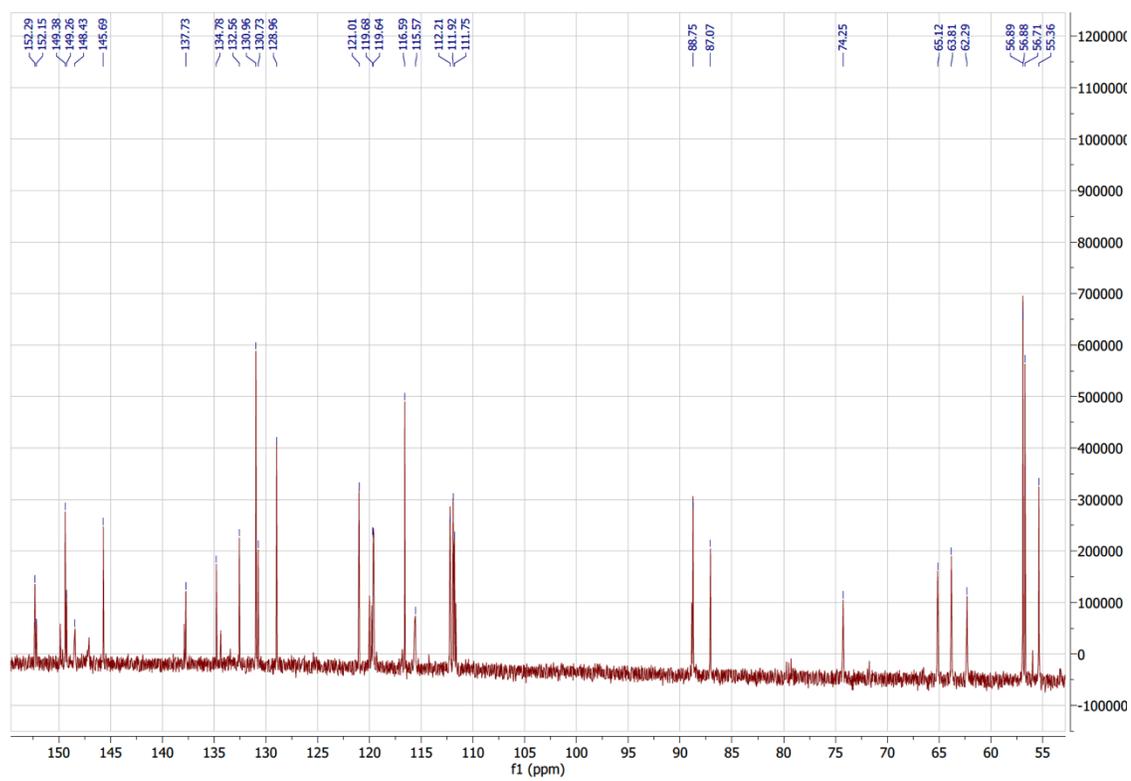
Figure 3.5 CNMR spectrum of G-βO4-G-β5-G trimer (compound 10).

The aromatic carbons (1, 2, 5, 6) in each monolignol appeared in regular region for aromatics at 110-140 ppm. Carbon 3 and 4 were deshielded due to the fact that they were bonded to an electronegative oxygen and shifted upfield (140-160 ppm). The carbons of the methoxy groups all appeared around 55 ppm. All gamma carbons (9, 9′, 9″), and the alpha and beta carbons on ring A (7, 8), and also the alpha carbon on ring B (7′) appeared about 20 ppm higher than the normal region for CH and CH2 (40-60 ppm) as they were all bonded to oxygen. The alpha and beta carbons on ring C (7″, 8″) were also observed in the regular vinyl region at around 130 ppm.
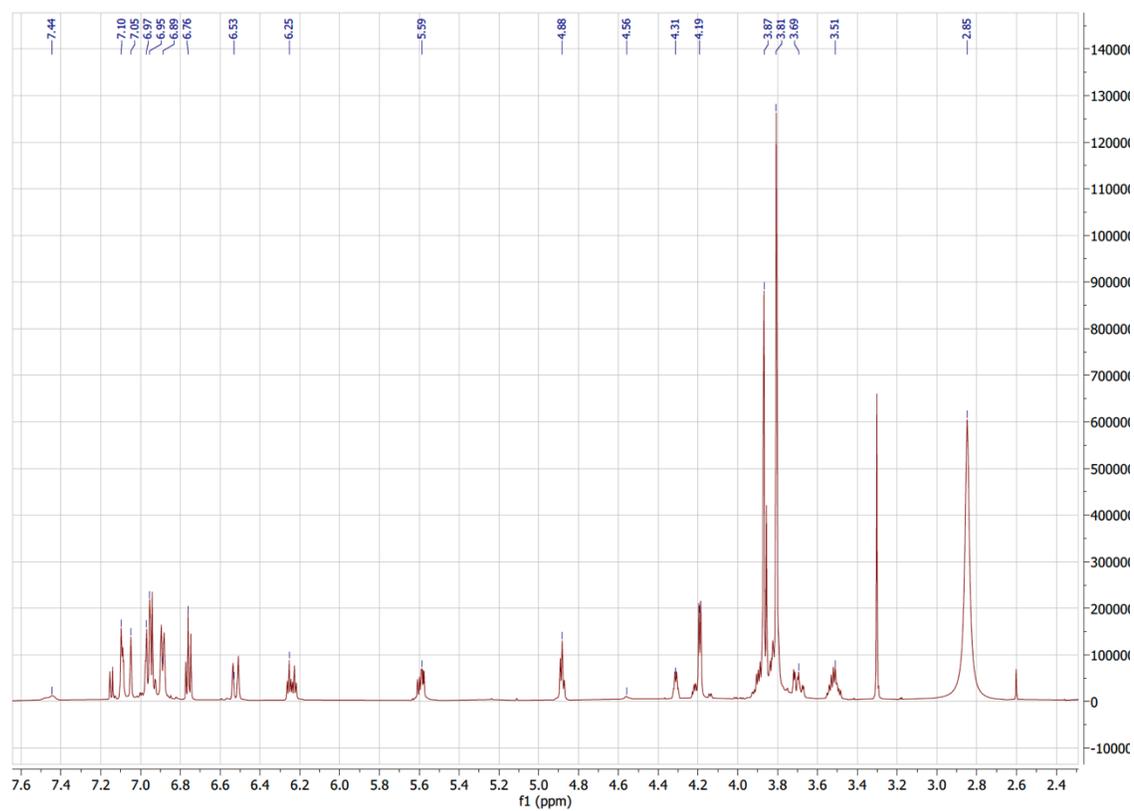
Figure 3.6 HNMR spectrum of G-βO4-G-β5-G trimer (compound 10)

All hydrogen atoms on the aromatic ring appeared at approximately 7 ppm. Similarly, all three methoxy hydrogens were observed as singlets at approximately 3.5 ppm, as expected. The hydrogens on the alpha and beta carbon on ring C (7″, 8″) were observed in vinyl region at 6-7 ppm and were coupled and splitting each other. The hydrogens on the beta carbon experienced splitting into triplet from two hydrogens on gamma and then into doublets by hydrogen on alpha and appeared as doublet of triplet. The reason for this is that the coupling constant, J, for vinyl hydrogens (16 Hz) is different from that of aliphatic hydrogens (6 Hz). Also, it seems like the double bond was in trans configuration, since the coupling constant in trans (11-18 Hz) is greater than cis (6-14 Hz) configuration. However, the other beta hydrogens on the other two rings appeared as quartets, which were splitting from three adjacent hydrogens at the same time.

Hydrogen on alpha and gamma on all three monolignols appeared as doublets which were splitting by hydrogens of beta. Similarly, the hydrogens on carbons 5 and 6 in rings A and B coupled and appeared as doublets.

It was observed that hydrogens on aliphatic hydroxyl groups were present at 2.5-4.5 ppm, and hydrogens on phenolic hydroxyl groups were present at 7.5 ppm, which is in the normal range for hydroxyl groups.
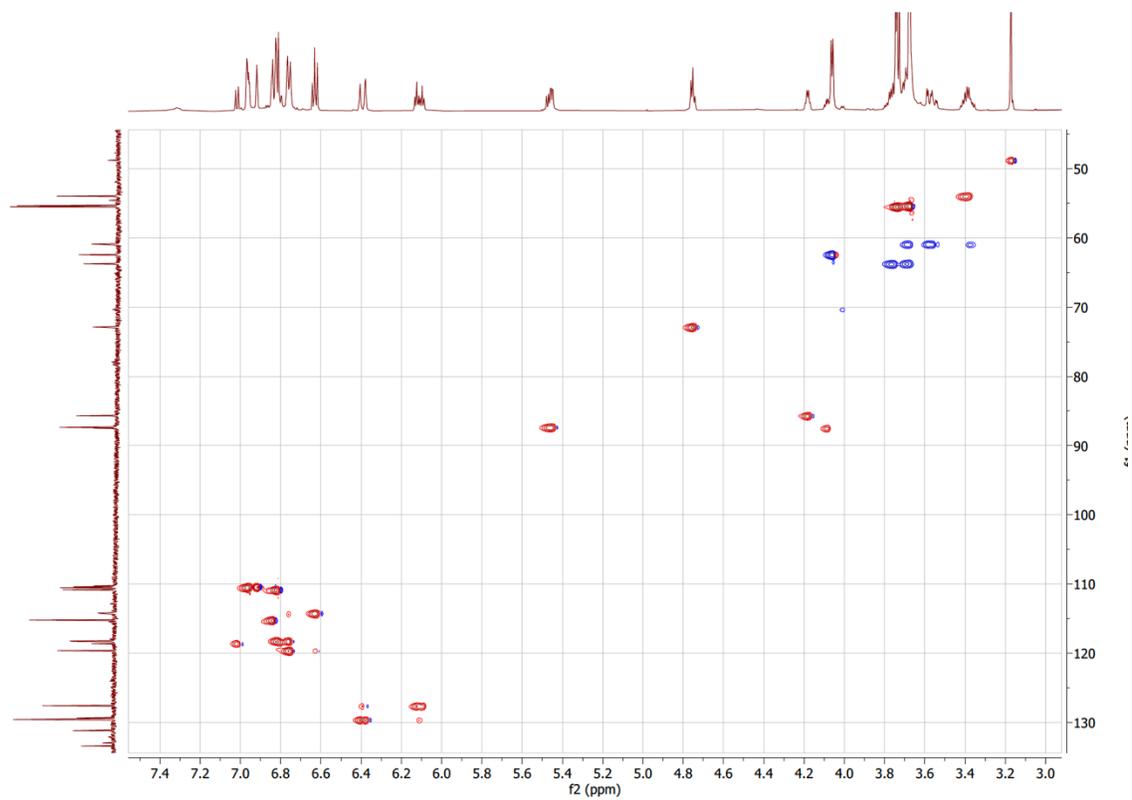
Figure 3.7 HSQC spectrum of G-βO4-G-β5-G trimer (compound 10)
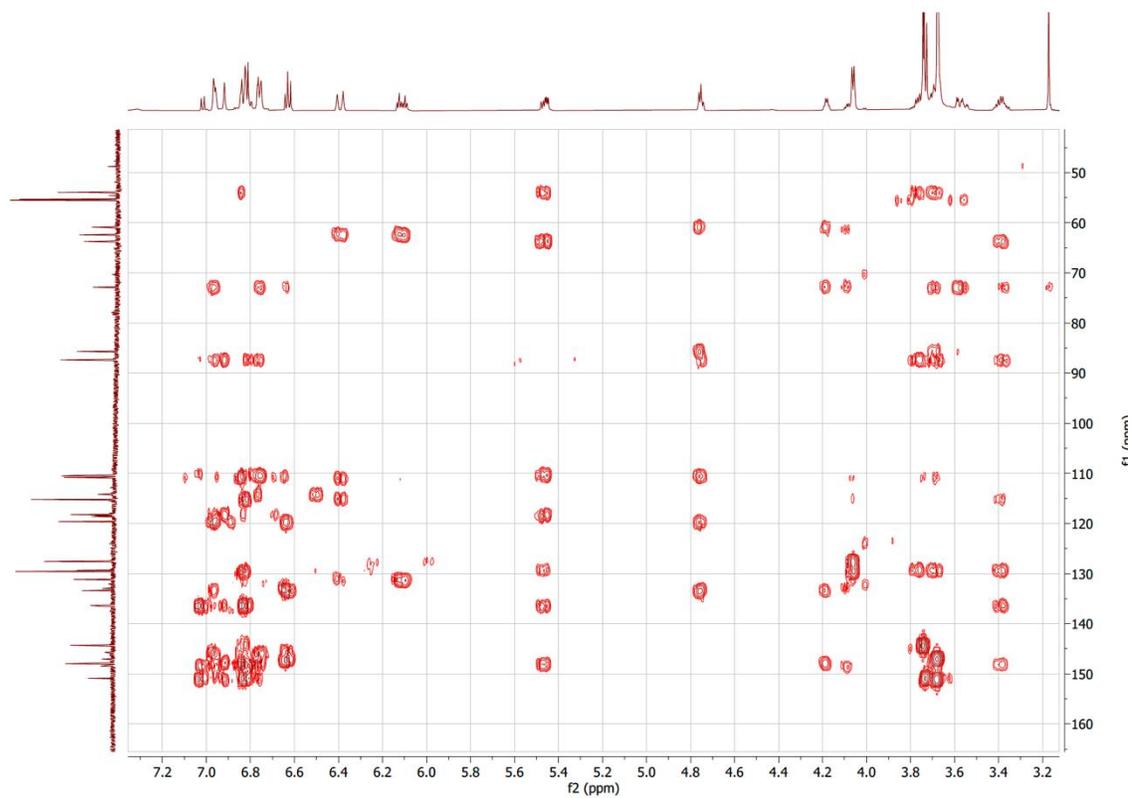
Figure 3.8 HMBC spectrum of G-βO4-G-β5-G trimer (compound 10)

As it was observed from the 2D carbon-hydrogen spectra, in HSQC all hydrogens exhibited correlation with the carbons to which they were bonded, as presented in Table 3.1.

In HMBC spectrum, the correlations between carbons and hydrogens of neighboring elements were observed and the linkage of the three monolignol was investigated.

The alpha and beta positions of ring A and ring B were the correlations that provided key information to unravel the bonding motif between each of the two monolignols. The alpha carbons on ring A and ring B (7, 7′) in CNMR appeared at 74.25 ppm and 88.75 ppm, respectively. Since 7′ was connected to an ether bond, it was slightly shifted upfield compared to 7 which was connected to hydroxy group. Therefore, the peak at 74.25 ppm in CNMR was for carbon 7. In HMBC, this carbon exhibited correlations with five adjacent hydrogens. The correlations at 4.31 ppm and 3.80 ppm were for hydrogens at beta and gamma carbons of ring A (8, 9). These correlations were also observed from carbons of beta and gamma (8, 9) to the hydrogen on carbon 7. In Figure 3.9, the double-sided arrows represent the correlations from both sides. The correlations from carbon 7 to hydrogens on 2, 5 and 6 provided the chemical shifts for the ring A as these positions were in the range of 3-carbon distance of the alpha carbon (7).

Carbon 8 and carbon 9 also exhibited correlation with each other. The other important correlation was the correlation of carbon 8 to carbon 4′ which proved the β-O-4 bond type between ring A and ring B.

The correlation between the carbons on ring A were also observed and presented in Figure 3.9.

Carbon 7′ and carbon 8′ also exhibited important correlations. Similar to carbon 7, on carbon 7′ the correlations between alpha, beta and gamma carbons were observed on both sides. Carbon 7′ also showed correlations to aromatic carbons on ring B, while the correlations between carbons in ring B indicated that they were bonded together. The important correlations on carbon 7′ were with aromatic carbons on ring C (4″ and 5″). The furan ring in the β-5 bond type, placed carbon 4″ and carbon 5″ in 2-atom distance from carbon 7′, making it possible to see the correlation from carbon 4″ and carbon 5″ with no hydrogen to the hydrogen on carbon 7′.

The correlation between carbon 8′ and three carbons on ring C (4″, 5″, 6″) also provided more information on the bond type between ring B and ring C.

Figure 3.9 displays all the correlations observed in the HMBC on the structure of the trimer. The double-sided arrows are the correlation from both carbon-hydrogen pairs, and it was observed that almost all adjacent carbons with hydrogen exhibited correlations with each other. The single sided arrows are likely for carbons that do not have hydrogen; thus, the correlation cannot be observed from the other side.
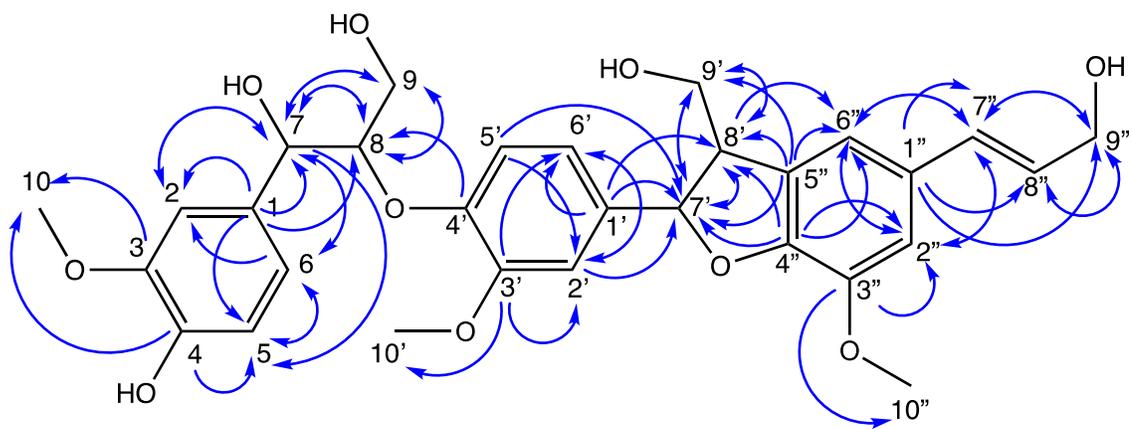
Figure 3.9 HMBC correlations on G-βO4-G-β5-G trimer (compound 10) structure.

3.3.2 Mass Spectrometry Analysis of the G-βO4-G-β5-G Trimer

The synthesized lignin trimer model compound was analyzed with mass spectrometry with adduct ionization that was well-studied by our research group. Lithium-adduct ions of lignin have been well-established in the previously published papers.

In the sequential analysis of macromolecules such as lignin, mass spectrometry has proven to be one of the most reliable methods. In previous work, our group was able to provide sequence-specific fragmentation of β-O-4 and β-β dimers and trimers with relatively stable fragments. Here, we analyzed a mix-linkage trimer with two of the most common linkages reported on natural lignin, which are β-O-4 and β-5 bonds. There is a high probability that these two bond types are adjacent in natural lignin, and fragmentation analysis of this trimer provides valuable information regarding lignin structural elucidation. A G-βO4-G-β5-G lignin trimer was analyzed with a Q-Exactive orbitrap mass spectrometer, and a detailed fragmentation analysis by HCD tandem mass spectrometry is provided.

The synthesized trimer contained three monomeric units which were all G units. If we consider the free phenol end of a lignin polymer as the phenol-terminus and the aliphatic hydroxyl as the aliphatic-terminus, in the synthesized model compound, the first two monomeric units, moving from the phenol-terminus to aliphatic-terminus, were linked by β-O-4 bond type and then with a β-5 bond type, it was linked to the last monomeric unit. The calculated exact mass of the trimer was found to be 554.2152 amu ($C_{30}H_{34}O_{10}$) and the lithiated molecular ion calculated to be m/z 561.2312 ([$C_{30}H_{34}O_{10}$]Li^+), which was observed by full scan on Q-Exactive (+) ESI mass spectrometer as m/z 561.2264

(error(ppm): -8.5526) (Figure 3.10). The peak observed at m/z 603.2109 is the trimer with the ion pair of LiCl ($[M+LiCl+Li]^+$) with a Li^+ ion, which commonly occurs during lithium adduct ionization and can be controlled by the concentration of the LiCl aqueous phase.

Another commonly observed peak is the mass-to-charge ratio of lithium bound dimer of the analyte. This peak is observed at m/z 1115.4370 ($[C_{60}H_{72}O_{20}]Li^+$) in the full scan (Figure 3.10).

The (+) HCD tandem mass spectrum of the trimer was obtained in a range of normalized collision energies (NCE). There was no fragmentation in 10 NCE, and the NCE was increased until lithiated fragment ions were observed. Therefore, the fragmentation analysis was conducted between 30 and 40 NCE.

Figure 3.10 Q-Exactive HRMS full scan of lithiated G-βO4-G-β5-G trimer (compound 10).

3.3.2.1   HCD Tandem Mass Spectrometry Experiment and Fragmentation of G-βO4-
         G-β5-G Trimer (Compound 10)

The fragmentation of the trimer showed a main cleavage of β-O-4 bond and multiple peripheral groups. The β-O-4 bond mainly cleaves at beta carbon and oxygen and this has been widely reported. The details of this cleavage have been previously reported by our research group in studies on various β-O-4 dimer and trimer models. Furthermore, this fragmentation pattern can be found in the literature for ESI ionization of β-O-4 model systems. Consistent with our trimer model study, the dominant peak appeared due to the cleavage of beta carbon and oxygen in the β-O-4 bond. This fragmentation is labeled as pathway a. As a result of this cleavage both fragments formed stable lithium adducts and were observed in the tandem mass spectrum. This fragmentation (pathway a) resulted in a G monomeric unit as m/z 203.0874 ($[C_{10}H_{12}O_4]Li^+$ calc. exact mass m/z 203.0890) which further fragmented and lost a hydroxy and form m/z 186.0851 ($[C_{10}H_{11}O_3]Li^+$ calc. exact mass m/z 186.0863, labeled as a1 fragmentation), and a lithiated β-5 dimer fragment at m/z 364.1466 ($[C_{20}H_{21}O_6]Li^+$, calc. exact mass m/z 364.1498). The β-5 fragment ion lost peripheral groups as water and formaldeyde with m/z 18 or m/z 30 or a combination of those two which were shown in fragmentation pathways a_2-a_7 in Figure 3.13.

Fragmentation on the trimer also indicated a loss of m/z 18 and m/z 30. This happened in a range of fragmentation of one m/z 18 or one m/z 30 up to the loss of two m/z 18 and two m/z 30 which resulted in fragment m/z 465.1859 ($[C_{28}H_{26}O_6]Li^+$, calc. exact mass m/z 465.1889). The proposed structures of these fragments are presented in Figure 3.13 as fragmentation pathways b and c.

The fragmentations for the loss of m/z 18 and m/z 30 were related to the loss of water or formaldehyde resulting from fragmentation of a hydroxy group on the alpha carbon of β-O-4 bond or the gamma carbon of β-5 bonding motif and formaldehyde from gamma carbon of either of the monomeric unit. Even though the structures are proposed to be fragmented in those pathways for peripheral groups, it is not necessarily in that order. It is a combination of multiple fragmentation on those positions to lose m/z 18 and/or m/z 30. These fragmentations provided valuable information about the peripheral groups and are critical to elucidate the structure of lignin oligomer.

To the best of my knowledge, no report has previously been published pertaining to the pathway d indicated in Figure 3.13. The cleavage of phenolic oxygen with the ring in β-O-4 moiety is not resulting in abundant stable fragment ions in lower NCEs. As the collision energy was increased the lithiated fragment ion intensities, resulting from this cleavage, increased to the point it become the base peak at NCE 40.

The fragmentation pathway d resulted in a monomeric unit of m/z 221.0979 ($[C_{10}H_{14}O_5]Li^+$, calc. exact mass m/z 221.1001) which further fragmented into a G monolignol, m/z 186.0851 ($[C_{10}H_{11}O_3]Li^+$, calc. exact mass m/z 186.0863) with pathways d_3 and a_1. The accurate mass unambiguously, proved formation of stable lithiated fragment ion for this cleavage. As a result of this fragmentation, the other part of the trimer should form lithiated ion at m/z 349.1627 ($[C_{20}H_{22}O_5]Li^+$). As the nominal mass was observed in HCD tandem mass at m/z 349.1234, interestingly, accurate mass data disprove formation of this fragment ion. The error in the obs. and calc. exact mass of the fragment was found to be -112.5550 ppm which is not in the acceptable range. The fragment peaks at m/z

331.1127 and m/z 319.1129 are clearly the loss of peripheral group with the fragmentation of m/z 18 and m/z 30, respectively.

High resolution mass spectrometry (HRMS) has emerged as a potent analytical tool in the characterization of complex molecules. Although the complete characterization of certain peaks remains challenging and necessitates further investigation, HRMS has demonstrated its efficacy in providing more precise molecular information. The fragment peak in question, for instance, could have been considered as the proposed fragment using lower resolution mass spectrometry such as quadrupole. However, with the high accuracy and resolution of HRMS, this assumption was dismissed. HRMS is capable of accurately measuring the mass-to-charge ratio of ions to several decimal places, making it possible to distinguish between closely related fragments and obtain a more comprehensive understanding of the molecular composition of complex samples.
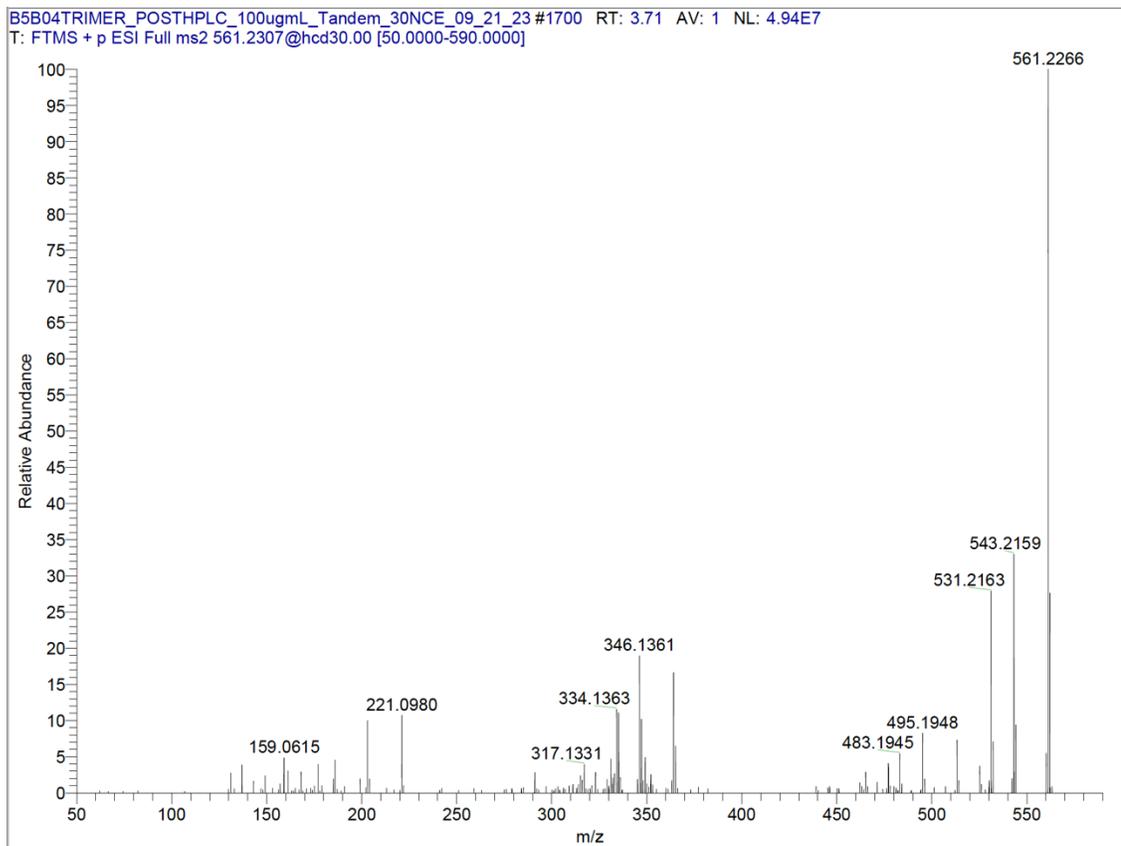
Figure 3.11 HCD tandem mass spectrum of G-βO4-G-β5-G trimer (compound 10) at 30 NCE.
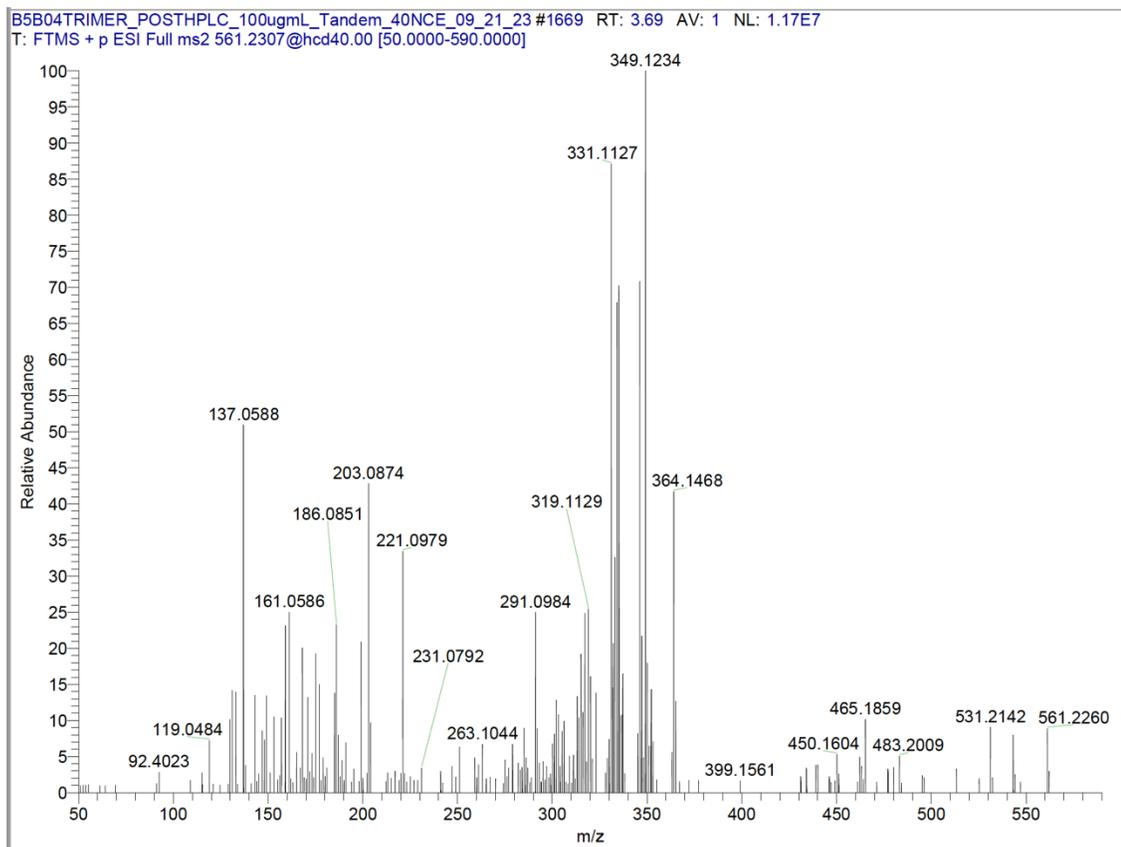
Figure 3.12 HCD tandem mass spectrum of G-βO4-G-β5-G trimer (compound 10) at 40 NCE.

Table 3.2 List of accurate mass data of fragment ions resulted from HCD analysis of lithiated G-βO4-G-β5-G trimer.(*) fragments which were disproved by HRMS.

Fragments	Obs. mass m/z	Calc. exact mass m/z	Error (ppm)
$[C_{10}H_{12}O_4]Li^+$	203.0874	203.0890	-7.8783
$[C_{10}H_{11}O_3]Li^+$	186.0851	186.0863	-6.4486
$[C_{20}H_{21}O_6^{.}]Li^+$	364.1466	364.1498	-8.7876
$[C_{28}H_{26}O_6]Li^+$	465.1859	465.1889	-6.4490
$[C_{10}H_{14}O_5]Li^+$	221.0979	221.1001	-9.9502
$[C_{20}H_{22}O_5]Li^+$	349.1234	349.1627	-112.5550*
$[C_{20}H_{20}O_4]Li^+$	331.1127	331.1522	-119.2805*
$[C_{19}H_{20}O_4]Li^+$	319.1129	319.1522	-123.1387*
$[C_{30}H_{32}O_9]Li^+$	543.2159	543.2206	-8.6521
$[C_{29}H_{30}O_8]Li^+$	513.2070	513.2101	-6.0404
$[C_{30}H_{30}O_8]Li^+$	525.2053	525.2101	-9.1392
$[C_{29}H_{28}O_7]Li^+$	495.1946	495.1995	-9.8950
$[C_{29}H_{32}O_9]Li^+$	531.2159	531.2206	-8.8475
$[C_{28}H_{28}O_7]Li^+$	483.1948	483.1995	-9.7268
$[C_{19}H_{19}O_5]Li^+$	334.1364	334.1393	-8.6790
$[C_{20}H_{19}O_5]Li^+$	346.1362	346.1393	-8.9559
$[C_{19}H_{18}O_4]Li^+$	317.1334	317.1365	-9.7750
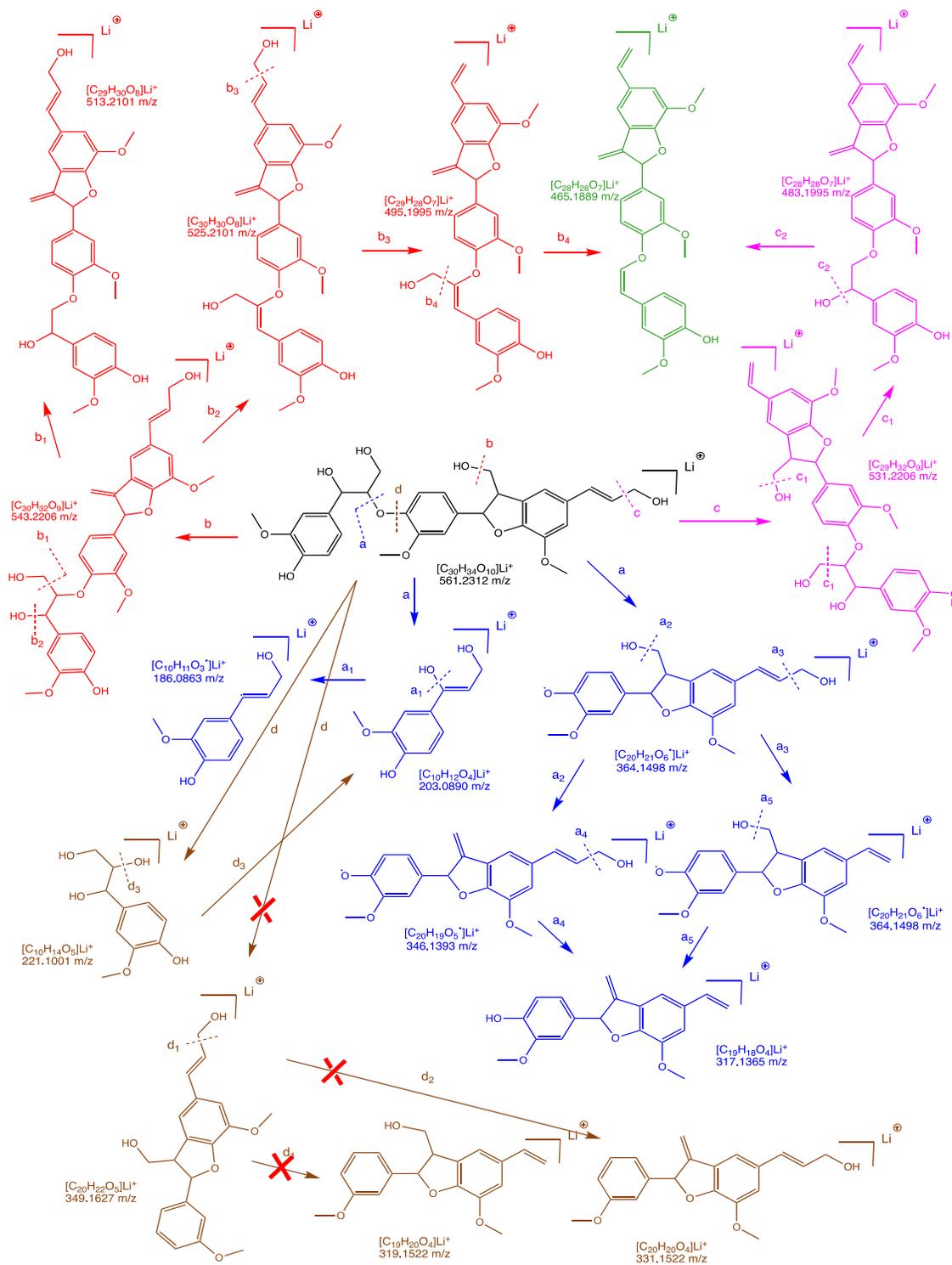
Figure 3.13 proposed fragmentation pathway for G-βO4-G-β5-G trimer (compound 10) and disproved fragments with HRMS.

3.4   Conclusion

In this study, a very important and valuable lignin oligomer containing two different bond types of β-O-4 and β-5 model compounds has been synthesized. The synthesis has been confirmed by various spectroscopic techniques such as HNMR, CNMR, HSQC and HMBC spectra. Furthermore, it has been proved and analyzed by high-resolution mass spectrometry (HRMS). The G-βO4-G-β5-G trimer synthesis accomplished by implementation of ferric chloride oxidative coupling to synthesize β-5 moiety and aldol coupling to synthesize β-O-4 moiety. The synthetic pathway was designed to retain the double bond on the aliphatic chain which plays an important role in lignin chemical behavior. The synthesized trimer purified with semipreparative HPLC and was subject to mass spectral analysis using lithium adduct ionization. The ions were introduced into the mass spectrometer using electrospray and subsequently analyzed using a Q-Exactive orbitrap mass spectrometer. The trimer formed stable lithium ion adduct and produced structurally informative fragments.

A detailed fragmentation pattern has been provided using Q-Exactive orbitrap HCD experiment by lithium cationization. Our study showed that the trimer with β-O-4 and β-5 bond type mainly undergoes fragmentation on β-O-4 moiety and loses peripheral groups. The loss of peripheral groups resulted in a decrease of m/z 18 and m/z 30, which were derived from the fragmentation of gamma hydroxy groups on all monomers or hydroxy group on alpha carbon in the β-O-4 bonding moiety. These fragments provide valuable information for structural elucidation of lignin oligomers and lignin breakdown products.

While this study successfully synthesized a lignin oligomer and provided detailed mass spectral analysis, there are some limitations that should be addressed. In this study, only β-O-4 moiety fragmentation was observed for the trimer and no fragmentation occurred on the β-5 moiety. Even though, detailed fragmentation on β-5 models is not reported in the literature, more investigation is required for understanding the pattern of fragmentation in this type of bonding. Additionally, the loss of peripheral group could not be identified with certainty as which monolignol gamma carbon is fragmented. This arises the limitation to fully elucidate the structure of oligomer. To address these issues future research could focus on isotopically labeled model compound and mass spectral analysis. The mass of isotopic labeled fragments can assist in defining fragments in a more unambiguous manner. This can also help unraveling fragmentation behavior of β-5 bond type models.

The results of this study have important implications for the field of lignin research, as the synthesis of this lignin oligomer provides a valuable tool for further investigation of lignin structure. Moreover, the detailed fragmentation pattern provided in this study can be used as a reference for the analysis of other lignin oligomers containing β-O-4 and β-5 bond types. Additionally, the methodology used in this study can be applied to the synthesis and analysis of other lignin oligomers, which may have different bond types and structures. Overall, this study contributes to a better understanding of lignin behaviour in mass spectral analysis and provides new opportunities for lignin structural elucidation in various lignin breakdown products.

CHAPTER 4. QUANTITATIVE DERIVATIZATION FOLLOWED BY REDUCTIVE
CLEAVAGE (QDFRC) ANALYSIS FOR EVALUATION OF BIOMASS
PRETREATMENTS

4.1   Introduction

Traditionally, lignin was considered as a waste stream from the pulping industry.

Pulping industries produce paper pulp from wood and other lignocellulosic biomasses[97-

99]. The process of pulping involves separating cellulose fibers from lignin and other non-

cellulosic components in wood and the resulting cellulose is then processed into paper

products. Lignin is the second most abundant natural polymer after cellulose. In the past,

lignin was typically burned as a fuel to generate energy for the pulping process, but

advances in technology have made it possible to extract and process lignin for other

uses[100]. For example, lignin can be used as a feedstock for the production of biomaterials,

bioplastics, and other high-value chemicals[101-103]. It can also be used as a natural adhesive

in wood products, such as plywood and particleboard[104, 105]. The lignin resulted from

pulping process is called kraft lignin. The chemicals used in the process of pulping, such

as sodium hydroxide and sodium sulfide alter the chemical properties of lignin and result

in a condensed kraft lignin. In recent years, alternative methods have been developed that

focus on lignin instead of cellulose extraction. These methods are generally called "lignin-

first pretreatments"[106].

4.1.1   Pretreatment

Pretreatment of biomass refers to the process of treating lignocellulosic biomass,

such as wood, agricultural residues, and energy crops, for conversion into fuels, chemicals,

and other value-added products. Biomass pretreatment is a critical step in the overall

production process, as it is necessary to break down the complex lignocellulosic structure of biomass and make its constituents more accessible to subsequent processing steps.

The primary goal of pretreatment is to separate three components of biomass: lignin, cellulose, and hemicellulose. There are many methods of biomass pretreatment, that have been proposed over the years. Some common pretreatment methods include chemical, physical, thermal, biological pretreatments[107, 108]. There has been reports of the use of combination of two different pretreatments to optimize the breakdown of lignocellulosic biomass[107]. Each of these pretreatment methods has its own benefits and drawbacks, depending on the specific biomass feedstock, conversion technology, and economic considerations.

As discussed in previous sections, traditional pretreatment procedures mainly focused on the extraction of cellulose and resulted in the alteration of chemical properties of remaining lignin. In recent years, lignin-first pretreatments have gained attention for their ability to preserve this valuable complex polymer and facilitate its valorization. The lignin-first pretreatment process involves the extraction of lignin from biomass using a solvent or other methods, followed by the separation of lignin from cellulose and hemicellulose components. The advantages of lignin-first pretreatment are that it can potentially reduce the cost and complexity of biomass pretreatment, as well as increase the value of the lignin fraction of the biomass[106]. Additionally, lignin is a high-energy-density material that can be used as a renewable feedstock for the production of biomaterials and biochemicals, making it an attractive option for sustainable production.

Pretreatment of biomass with organosolv is considered to be one of the most promising methods for delignification of biomass. Resulting in relatively clean cellulose

solid, aqueous stabilized hemicellulose and solid precipitated lignin. In spite of this, the high temperature and energy required to operate this method prevent its widespread use in the industrial sector[41, 109-113]. Hydrotropic pretreatment uses para-toluene sulfonic acid (p-TsOH) to solubilize lignin. Since this pretreatment involves lower temperatures and energy consumption, as well as the fact that the acid can be reused, it constitutes a sustainable process[39, 114]. Acetic/formic acid pretreatment developed by the Compagnie Industrielle de la Matière Végétale (CIMV) is based on treatment of biomass with a mixture of acetic acid, formic acid and water. This pretreatment was also reported to produce more valuable lignin together with pure cellulose for paper industries[40].

In this study five different pretreatments ethanosolv, dioxosolv, Co-solvent Enhanced Lignocellulosic Fractionation (CELF), acetic/formic acid, and hydrotropic pretreatments were employed as a lignin etraction method and were evaluated for the quality of extracted lignin using qDFRC analytical technique followed by GC-MS analytical analysis.

4.1.2   Quantitative Derivatization Followed by Reductive Cleavage (qDFRC)

Derivatization Followed by Reductive Cleavage (DFRC) is one of the most common degradative methods for the analysis of lignin next to thioacidolysis. As a result of this analytical method, β-aryl-ether bonds, commonly known as β-O-4, are selectively cleaved[115]. In the derivatization step the alpha carbon is brominated by acetyl bromide in acetic acid. It was also observed that the gamma carbon is acylated in this step. The reduction of β-bromo-ethers is well studied in organic chemistry[116]. In this case, zinc dust was used to cleave the ether bond and form a double bond between alpha and beta carbons,

which is the original structure of the monolignols. As a result of the DFRC method, the monolignol (acylated) is the final product. This makes DFRC a unique lignin deconstructive method to breakdown lignin to its constituent monolignols from natural lignin[116].

As mentioned earlier, the most abundant bonding type in lignin is β-O-4 and DFRC method selectively cleaves this bond type. As a result, it is possible to gain an understanding of the amount of lignin in biomass by quantifying the retrieved monolignol. Nevertheless, this may not apply to lignin that has been extracted for biomass by pretreatment methods. During the process of extracting lignin, condensation can pose a significant challenge[117]. When lignin is liberated from the other component of the biomass, it can cross-link and form C-C bonds. This process can make lignin, this complex polymer, even harder to breakdown because the β-ether bonds are not accessible to the DFRC reagents. Thus, a lower yield would be observed for monolignols. Lignin condensation can also produce challenges for valorization of lignin. In this study, deuterium labeled internal standard were used to quantify the amount of liberated monolignols. These data were compared with the same value for untreated biomass to evaluate the efficiency of the pretreatments.
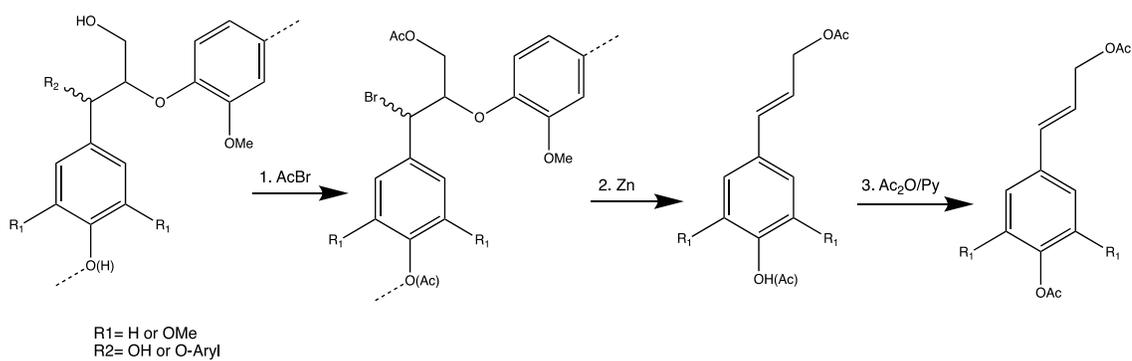
Figure 4.1 The chemical transformations associated with DFRC method[116].

4.2   Materials and Methods

4.2.1   Chemicals

All reagents were used without further purification. Acetyl bromide, acetic acid, dioxane, pyridine, ethyl acetate, sodium sulfate, hexane, hydrochloric acid, para-toluene sulfonic acid (p-TsOH), sulfuric acid, formic acid, and chloroform were purchased from Fisher Scientific (Fair Lawn, NJ, USA). Zinc dust, acetic anhydride, d_6-acetic anhydride, silica gel, and tetrahydrofuran (THF) were purchased from Sigma-Aldrich (Saint Louis, MO, USA). Ammonium chloride, ethanol, and potassium carbonate were purchased from Alfa Aesar (Ward Hill, MA, USA). Dichloromethane was purchased from VWR (Radnor, PA, USA). Wheat straw biomass (2mm) was obtained from University of Kentucky, Dr. Montross lab in Biosystems and Agricultural Engineering department (Lexington, KY, USA).

4.2.2   Methods

4.2.2.1   Preparation of Deuterium Labeled and Regular Acylated Monolignols

All three types of monolignols were prepared as described in Section 3.2.2.2.2. In the preparation of H, G, and S monolignols, p-coumaric acid, ferulic acid, and sinapic acid were used as starting materials, respectively. Each monolignol (100 mg) was dissolved in acetone (3 mL). Deuterium labeled acetic anhydride (~260 μL, 1:5 molar ratio) and potassium carbonate (~380 mg, 1:5 molar ratio) were added. The solution kept at 40°C and stirred overnight. The solution was separated from the solid potassium carbonate and extracted with ethyl acetate (5 mL × 3 times) and saturated ammonium chloride (5 mL).

116

The collective organic layer was dried under reduced pressure. The deuterium labeled acylated monolignols were purified using silica gel gravimetry column(~98% yield). For the regular acylated monolignols, regular acetic anhydride was used.

Table 4.1 GC-MS observed m/z for synthesized acylated monolignol and their main ion peak fragment m/z.

	Calculated m/z for $M^{+\bullet}$	Observed m/z	Base peak fragment m/z
D_6-acylated H monolignol	240	240	196
H-acylated H monolignol	234	234	192
D_6-acylated G monolignol	270	270	226
H-acylated G monolignol	264	264	222
D_6-acylated S monolignol	300	300	256
H-acylated S monolignol	294	294	252

4.2.2.2   Calibration Curve

Calibration curve for all three monolignols was obtained using 0.5 mg/mL concentration for deuterium labeled acylated monolignol as internal standard and varying range of concentration as 0.1, 0.2, 0.5 and 1 mg/mL of regular acylated monolignol in chloroform. The lowest concentration sample had 0.1 mg/mL concertation of regular acylated monolignol and 0.5 mg/mL concentration of deuterium labeled acylated monolignol and the highest concentration had 1 mg/mL concentration of regular acylated monolignol and 0.5 mg/mL concentration of deuterium labeled acylated monolignol. Each sample prepared with 4 replicates and injected on an Agilent (Santa Clara, CA, USA) GC-MS. The extracted ion chromatogram of the most abundant ion peak for the regular and deuterium labeled acylated monolignol was obtained as listed in Table 4.1 (192 and 196 for H monolignol, 222 and 226 for G monolignol and 252 and 256 for S monolignol). The integrated peak areas were used to draw the calibration curve. X axis represents concentrations of acylated monolignol. Y axis represents the integrated peak area of regular acylated monolignol divided by that of deuterium labeled acylated monolignol.

4.2.2.3   Quantitative Derivatization Followed by Reductive Cleavage (qDFRC)

The qDFRC experiment was followed as described in Lu&Ralph (2015) with some modification[116]. The qDFRC was used for biomass and extracted lignin. Biomass (5 mg) or extracted lignin (2 mg) was added to 20% (v:v) acetyl bromide in acetic acid solution (2.5 mL) with a magnetic stirrer. The mixture stirred at 50°C for 2.5 h. The solvent was removed under nitrogen in a heating block set at 25°C. After removing the solvent, ethanol (2 mL) was added and vortexed and the dried under nitrogen to remove any residual acetyl

bromide. A mixture of dioxane:acetic acid:water (5:4:1, v:v:v) (2.5 mL) and zinc dust (50 mg) were added to the dried residue and stirred in room temperature for 15 minutes. The solution then extracted with ethyl acetate (5 mL × 3 times) and saturated ammonium chloride (5 mL) with a pH of 2 (adjusted by adding 1N HCl). The organic layer dried under nitrogen. A 1:1 solution of acetic anhydride and pyridine (5 mL) was added to the residue and stirred for an hour. The solvent was removed under nitrogen with heating block set at 25°C. The residue was subject to solid phase extraction (SPE) with a 3 mL silica cartridge (500 mg per column, J.T.Baker, S. A. Poland). The cartridge was conditioned with hexane (3 mL). The residue was loaded on the cartridge with methylene chloride (100 µL) and eluted with mixture of 1:5 (v:v) ethyl acetate:hexane (12 mL). The solvent was evaporated under nitrogen and the residue was resuspended in internal standard stock solution (100 µL). The stock solution of internal standard contains all three deuterium labeled acylated monolignols with the concentration of 0.5 mg/mL.

4.2.2.4   Pretreatments

All pretreatment procedures were conducted by the optimal literature reported conditions, details below.

4.2.2.4.1   ETHANOSOLV PRETREATMENT

Ethanosolv pretreatment was conducted with 1:7 (w:v) ratio of biomass to 60% aqueous ethanol solution. Wheat straw biomass (8 g) was added to 60% aqueous ethanol solution (56 mL) using 1.25%wt of H2SO4 (1.25%wt to water) as acid catalyst. The

pretreatment was conducted in a 4842 Parr reactor (Parr Instrument Company, Moline, IL, USA) at 160°C for 60 minutes. After cooling the solution to ~70°C, the solid was filtered (Fisherbrand filter paper, qualitative: P8, porosity coarse flow rate: fast, Pittsburg, PA, USA) and washed with 60% aqueous ethanol. The filtrate then slowly was added to 600 mL DI water and left overnight for lignin precipitation. The clear solution on top decanted and lignin recovered by centrifugation and dried in room temperature[109, 111, 112].

4.2.2.4.2 DIOXOSOLV PRETREATMENT

Dioxosolv pretreatment was conducted with 1:12 (w:v) ratio of biomass to aqueous dioxane. Wheat straw biomass (2 g) was added to 9:1 (v:v) dioxane:water mixture (24 mL) with 0.2 M HCl (to water) as acid catalyst. Reflux was performed under nitrogen at 87.6°C for 3 hours. The solid cellulose and hemicellulose were filtered (Fisherbrand filter paper, qualitative: P8, porosity coarse flow rate: fast, Pittsburg, PA, USA) after cooling the solution and rinsed with 9:1 (v:v) dioxane:water 3 times. The resulting filtrate was added slowly to ~200 mL of acidified water with pH 2 (adjusted with 1 N HCl) while stirring for 20 min. After leaving the suspension overnight, lignin precipitated from the suspension which centrifuged and dried in room temperature[41].

4.2.2.4.3 CO-SOLVENT ENHANCED LIGNOCELLULOSIC FRACTIONATION (CELF)

Co-solvent Enhanced Lignocellulosic Fractionation (CELF) was conducted with 7.5%wt of biomass to 1:1 (v:v) mixture of THF:water. Wheat straw biomass (4 g) was added to 1:1 (v:v) THF:water mixture (56 mL) with 0.5%w sulfuric acid (to water) as catalyst. The pretreatment was conducted in a 4842 Parr reactor (Parr Instrument

Company, Moline, IL, USA) at 150°C for 30 minutes. After cooling down the mixture to ~70°C the solid filtered (Fisherbrand filter paper, qualitative: P8, porosity coarse flow rate: fast, Pittsburg, PA, USA) with vacuum filtration. The solution was neutralized with ammonium hydroxide (pH ~7). The THF was removed under reduced pressure in room temperature. The resulting mixture of precipitated lignin and water dried overnight at room temperature[118-120].

4.2.2.4.4 ACETIC/FORMIC ACID PRETREATMENT

Wheat straw biomass (4 g) was added to a mixture of acetic acid: formic acid: water (5:3:2 v:v:v) (40 mL) with 1:10 (w:v) ratio. After heating the mixture to 60°C for an hour, the mixture was refluxed for three hours at 107°C. The solid was filtered (Fisherbrand filter paper, qualitative: P8, porosity coarse flow rate: fast, Pittsburg, PA, USA) by vacuum filtration and the lignin was precipitated by concentration of the filtrate and adding to 400 mL water (400 mL). Centrifugation and drying in room temperature resulted in the collection of precipitated lignin[40].

4.2.2.4.5 HYDROTROPIC PREREATMENT

Wheat straw biomass (4 g) was added to preheated 60%wt aqueous p-toluene sulfonic acid (p-TsOH) solution (40mL) at 80°C with 1:10 w:v ratio. The mixture kept at 80°C with stirring for 30 min. The solid immediately filtered (Fisherbrand filter paper, qualitative: P8, porosity coarse flow rate: fast, Pittsburg, PA, USA) by vacuum filtration and the filtrate immediately diluted to 3%wt of p-TsOH. The solution left overnight for

lignin precipitation, and lignin was extracted by centrifugation and drying in room temperature[39, 114, 121].

4.2.3 Gas Chromatography-Mass Spectrometry

The GC-MS analysis was conducted on an Agilent 5973 MSD equipped with HP 6890 GC and HP 7683 injector controlled by ChemStation D.03.00611. A DB-5HT (Agilent Santa Clara, CA, USA). A GC column with 15m length, internal diameter of 250 μm and film thickness of 0.1μm was used. The injector temperature was set to 250 °C. The oven temperature was set at 100 °C and held for 3 minutes, then ramped at 15 °C/min to 280 °C and held for 10 minutes for a total method time of 25 minutes. An injection split ratio of 50:1 was used for analysis.

4.3 Discussion

4.3.1 Calibration Curve and Internal Standard

The lignin breakdown products from DFRC are acylated monolignols. In order to quantify the amount of retrieved monolignol, a calibration curve with known concentration is required. Isotopically labeled internal standards are widely used for quantification methods[122]. The internal standard used in this work is deuterium labeled acylated monolignol. The deuterium labeled acetyl group contain three deuterium atoms and each monolignol was acylated on the phenolic and aliphatic gamma hydroxyl groups resulting in an addition of 6 amu to the molecular weight of the monolignol. Reconstructed ion chromatograms were used on GC-MS to resolve the close retention time of the analyte and

the internal standard. The most abundant ion peak for each standard is selected for reconstructed ion chromatogram (Table 4.1).

The concentration for the H-acylated monolignols in the standard curve were ranged from 0.1, 0.2, 0.5 and 1 mg/mL to maintain the linear response of the GC-MS. The concertation for deuterium labeled acylated monolignol was kept constant at 0.5 mg/mL. The integrated peak area of each of the monolignols recorded and the calibration curve obtained with x axis as the concentration of H-acylated monolignol and y axis as the ratio of the integrated peak area of the H-acylated monolignol over the deuterium labeled acylated monolignol. Each sample prepared with 4 replicates and averaged to graph the calibration curve.

In the last step of the qDFRC the final residue was resuspended in 100 μL of the internal standard with the concentration of 0.5 mg/mL to all the monolignols.
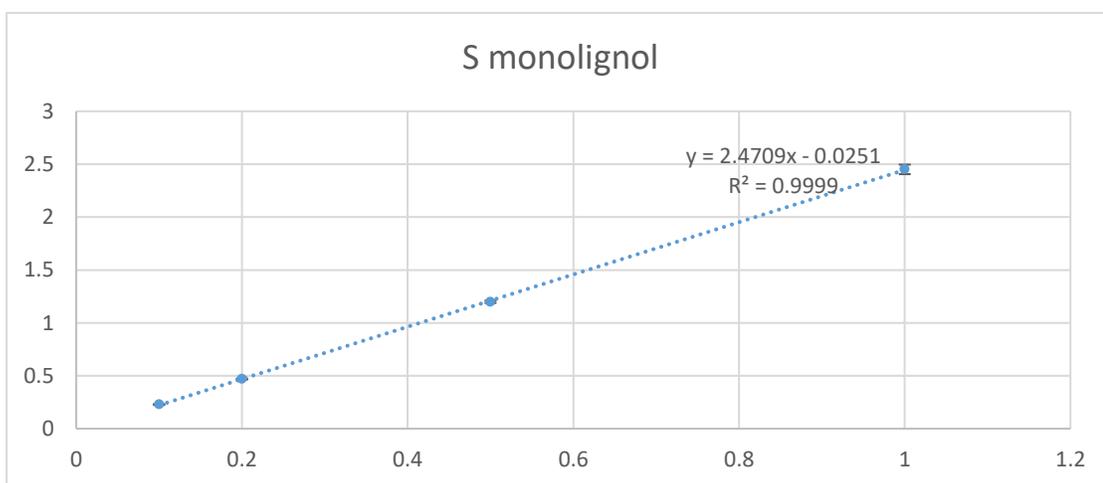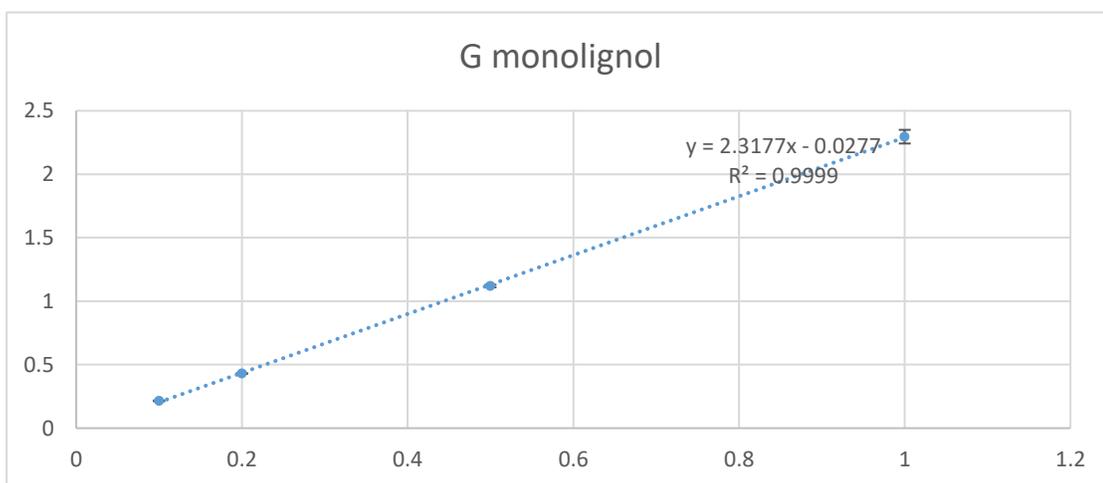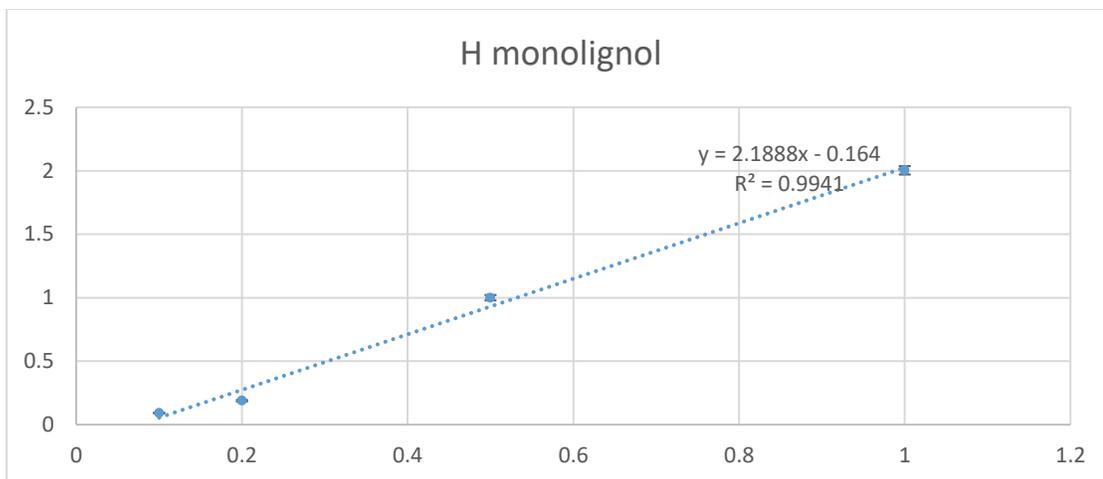
Figure 4.2 Calibration curve for GC-MS response of acylated monolignols

4.3.2   Quantitative Derivatization Followed by Reductive Cleavage (qDFRC)

The quantitative Derivatization Followed by Reductive Cleavage (qDFRC) was applied to evaluate different pretreatment and the quality of extracted lignin. The regular DFRC can provide only the ratio of S and G monolignol however this ratio might remain constant even if processes such as condensation occur on lignin. The S/G ratio does not reveal any information of the quantity of the monolignol liberated by the breakdown method. It is possible, for instance, that in a sample of extracted lignin, only a fraction of β-O-4 bonds were broken and most of these bonds were not accessible, however the S/G ratio could remain constant. With qDFRC, the amount of monolignol is quantified and compared to the similar value in biomass and other pretreatments. This data can provide us with important information regarding the severity of the pretreatment, the extent of condensation, and the overall quality of the lignin.

Applying this method, five different pretreatments were evaluated by comparing the amount of liberated monolignol in each pretreatment and untreated biomass with wheat straw. A detailed analysis is provided regarding the extent of lignin recovered from biomass as well as a quantitative and qualitative evaluation.

4.3.3   Pretreatments

Five pretreatments selected for this study were dioxosolv, ethanosolv, co-solvent lignocellulosic fractionation (CELF), acetic/formic acid and hydrotropic. The first three are generally categorized as organosolv pretreatment. In essence, organosolv pretreatment involves treating biomass with an aqueous organic solvent and small quantity of acid in order to facilitate the dissolution of the lignin component of the biomass. This step is

conducted in varying time and temperature. After filtering out undissolved cellulose and hemicellulose, lignin will be extracted by adding an antisolvent, which is primarily water.

Hydrotropic pretreatment involves the use of a small amount of a hydrotropic agent to enhance the solubility of lignocellulosic biomass in water. Hydrotropic agents are typically organic compounds, such as p-toluene sulfonic acid, that can increase the solubility of hydrophobic molecules, such as lignin, by forming a micellar solution, which can reduce the need for high temperature and pressure.

Acetic/formic acid pretreatments, also work with the same concept of dissolution of lignin in an organic acid solution at varying time and temperature, separation of solution from solid by filtration, and extracting the lignin by adding antisolvent.

As a means of validating the results from pretreatments, nine data point were obtained for each method. Wheat straw biomass was treated with each method three times. Also, the qDFRC analysis were conducted three times on each extracted lignin from the previous replicates, resulting nine separate data point for each pretreatment. An ANOVA test was taken from these three groups to prove there are no significant difference between the groups and consider all nine data point for each pretreatment.

4.3.4   Gravimetric Analysis

In this study, various masses of wheat straw biomass were used for pretreatments, and the amount of lignin extracted is presented as a percentage by mass of the biomass weight before pretreatment. The amount of lignin reported in the literature for wheat straw is between 11%-26%[123]. The range of extracted lignin with these five pretreatments is 1.19%-12.25% with dioxosolv as the lowest yield and hydrotropic as the highest yield of

lignin. The average percentage of the extracted lignin from each pretreatment is presented in Table 4.2. Hydrotropic and ethanosolv showed to be the most effective methods in extracting lignin as they obtained 12.25% and 11.97% lignin from wheat straw biomass, respectively. CELF is the next pretreatment in the trend with 8.02% extracted lignin followed by acetic/formic acid with 6.04%. The least amount of extracted lignin was observed in dioxosolv pretreatment with 1.19%.

As it is important to extract all possible lignin from the biomass, the quality of the extracted lignin is also important. When lignin is isolated from hemicellulose and cellulose it is potent to form C-C bonds. By forming these new, harder to break C-C bonds, lignin become more resilient to most of lignin breakdown methods. The process of forming these new C-C bonds and more cross-linking of lignin polymer is commonly known as lignin condensation.

Condensation of lignin is one of the common artifacts that can be observed during the process of pretreatment. The extent of condensation may vary depending on the severity of the pretreatment. Therefore, the quality of extracted lignin from each of these pretreatments was investigated.

Table 4.2 Recovered lignin percentages for all pretreatments. (*) three replicates were conducted for each pretreatment, but one replicate did not yield usable data due to technical difficulties.

Pretreatment	N	Recovered Lignin wt%
Dioxosolv	3	1.19%±0.05%
Ethanosolv	3	11.97%±0.31%
CELF	3	8.02%±0.86%
Acetic/formic acid	2*	6.04%±1.15%
Hydrotropic	3	12.25%±1.15%

4.3.5 Qualitative Analysis

Traditionally, DFRC was used to find S/G ratio. After treating biomass or extracted lignin with DFRC, the resulting residue is analyzed with GC-MS[115]. The integrated areas of G and S provided information on S and G monolignol content of the biomass. The purpose of discussing the qualitative analysis of the extracted lignin with these pretreatments is to compare the change in S and G content of the extracted lignin with untreated biomass. It is also worth noting that by comparing quantitative results with qualitative S/G ratios, even though the ratio can provide information on which monolignol is more condensed, it does not provide information on the overall condensation.

The reduction in observed monolignols can result from multiple sources. As previously explained, in the process of DFRC, the alpha hydroxy was brominated in the acetyl bromide digestion step of the biomass, resulting in the formation of a double bond after the reductive cleavage step and observation of monolignols. One possibility for the reduction of observed monolignols could be due to the structural alteration of lignin during the pretreatment process, resulting in the elimination of alpha hydroxy. This alteration prevents the bromination of the alpha carbon, resulting in fewer observed monolignols. Another possibility could be the cleavage of some β-O-4 bond types in the pretreatment process, resulting in a reduction of observed monolignols. It is also possible that the reduction of monolignols resulted from condensation on the lignin during pretreatment, which was discussed earlier. Although we cannot reject other possibilities, condensation was considered the main reason for the difference in the results for observed monolignols from qDFRC due to its high reporting in literature and for the simplification of the interpretation of results.

Table 4.3 shows the qualitative analysis of DFRC results from extracted lignin by pretreatments and untreated biomass. Values for H, G and S monolignol are percentage of the integrated peak area of each monolignol (most abundant ion peak) over the total integrated area of all three monolignol. In the last column, the S/G ratio is also provided.

Hydrotropic pretreatment which has the highest yield for lignin extraction have an increased S/G ratio (1.1) compared to untreated biomass (0.5). Which shows more condensation on G monolignol. There was no significant change for H monolignol as the percentage decreased from 2.4 to 2.2 compared to untreated biomass.

In ethanosolv and CELF also, there was an increase in S/G ratio (0.9 and 0.7 respectively), which interpreted as more condensation on G monolignol. As ethanosolv have almost the same percentage for H monolignol, compared to untreated biomass, this value is increased in CELF pretreatment which means both G and S monolignol were less accessible for DFRC cleavage compared to untreated biomass. However, increases S/G ratio for CELF shows that condensation was more intense on G monolignol.

Acetic/formic acid and dioxosolv pretreatments had a slight decrease on S/G ratio and increase on H monolignol. Which means condensation on both G and S monolignol but more on S monolignol.

Table 4.3 Percentages of each monolignol observed in GC-MS without including internal standard and S/G ratios. (*) The data points which had significant difference excluded from the group and discussed in Section 4.3.7.

Pretreatment	N*	Percentage			S/G
		H monolignol	G monolignol	S monolignol	
Dioxosolv	9	5.1%±0.5%	67.3%±1.4%	27.6%±1.5%	0.4±0.1
Ethanosolv	6	2.2%±0.4%	51.5%±1.9%	46.4%±1.8%	0.9±0.1
CELF	9	3.0%±0.1%	57.8%±1.7%	39.2%±1.8%	0.7±0.1
Acetic/formic acid	6	2.9%±0.5%	68.4%±1.6%	28.7%±1.9%	0.4±0.1
Hydrotropic	6	2.2%±0.3%	48.1%±6.2%	49.7%±6.4%	1.1±0.3
Untreated biomass	3	2.4%±0.1%	63.7%±4.8%	33.9%±5.0%	0.5±0.1

4.3.6   Quantitative Analysis

In previous section, with qualitative analysis, the proportional change in H, G and S monolignol was investigated. In quantitative analysis, each monolignol is quantified and compared to the same value driven from untreated biomass. The lignin content of wheat straw biomass is reported to be 11-26 in the literature[123]. The composition of wheat straw biomass might be affected by specific cultivator, soil type, fertilizer treatment and other growth condition[123]. The highest yield for lignin among the studied pretreatments was 12.25%.

Quantitative analysis in qDFRC, uses the internal standard to find the mass of each monolignol. These masses then divided by the initial mass of lignin or biomass to report a ratio of the liberated monolignols per mass of lignin or biomass. Therefore, to compare these values the denominator of these ratios should be the same. The value obtained for untreated biomass was per mg of biomass however this value for the extracted lignin is per lignin mass. By assuming that 12.25% of the wheat straw biomass is lignin, as it was the highest yield of extracted lignin among the five pretreatments, which is consistent with the literature reported value[123], we can use a conversion factor to report the values for untreated biomass as mass of each monolignol per lignin in untreated biomass. The comparison of these values for monolignols per lignin in pretreated and untreated biomass will provide valuable information regarding the extent of condensation.

The amount of liberated monolignol in all pretreatments was at least three times less that what it was observed in untreated biomass. According to this observation, monolignols, or in more general term, lignin was less susceptible of cleavage on β-aryl-

ether bonds after it has been processed with different pretreatment. Table 4.4 and Figure

4.3 show qDFRC results of extracted lignin from pretreatments and untreated biomass.

Table 4.4 qDFRC results for pretreatments and untreated biomass. (*) values for untreated biomass are adjusted with the assumption of 12.25% lignin in biomass which was driven from the highest yield of recovered lignin. The values are per mg of lignin in biomass.

Pretreatments	H monolignol (µg/mg)	G monolignol (µg/mg)	S monolignol (µg/mg)
Dioxosolv	6.6±0.5	33.0±3.1	10.5±1.1
Ethanosolv	4.0±0.4	15.3±1.0	12.1±1.1
CELF	4.9±0.7	30.7±4.5	18.2±3.7
Acetic/formic acid	4.6±0.4	19.7±3.9	6.3±2.3
Hydrotropic	3.6±0.5	4.3±0.6	3.8±0.7
Untreated biomass*	17.0±0.2	121.0±13.6	53.7±8.4
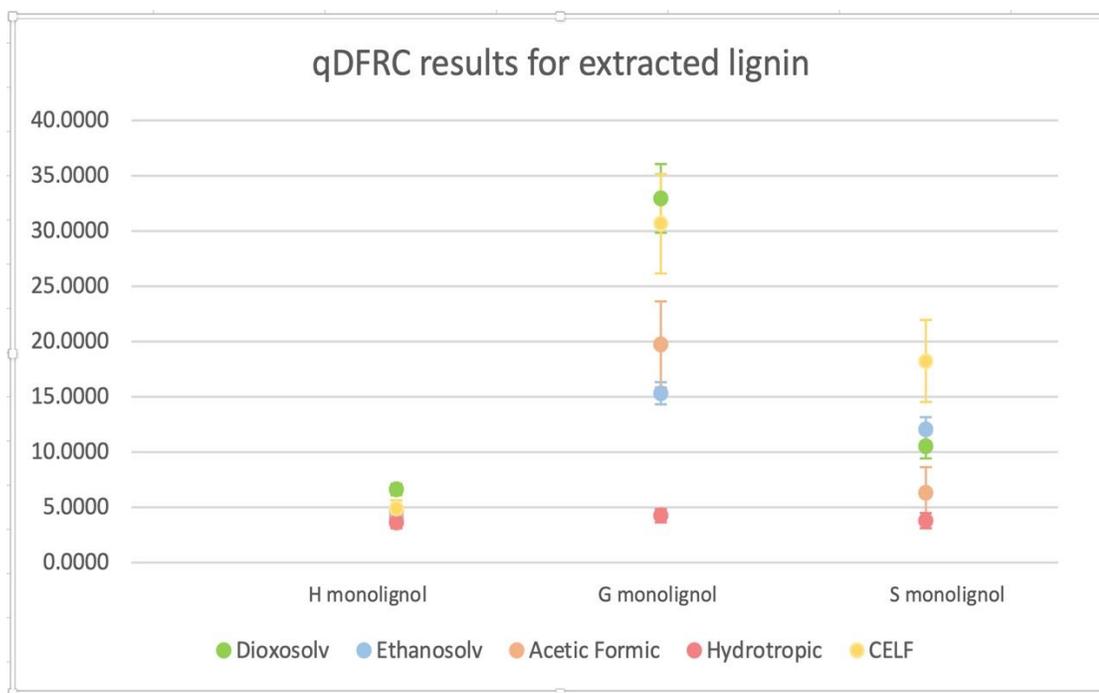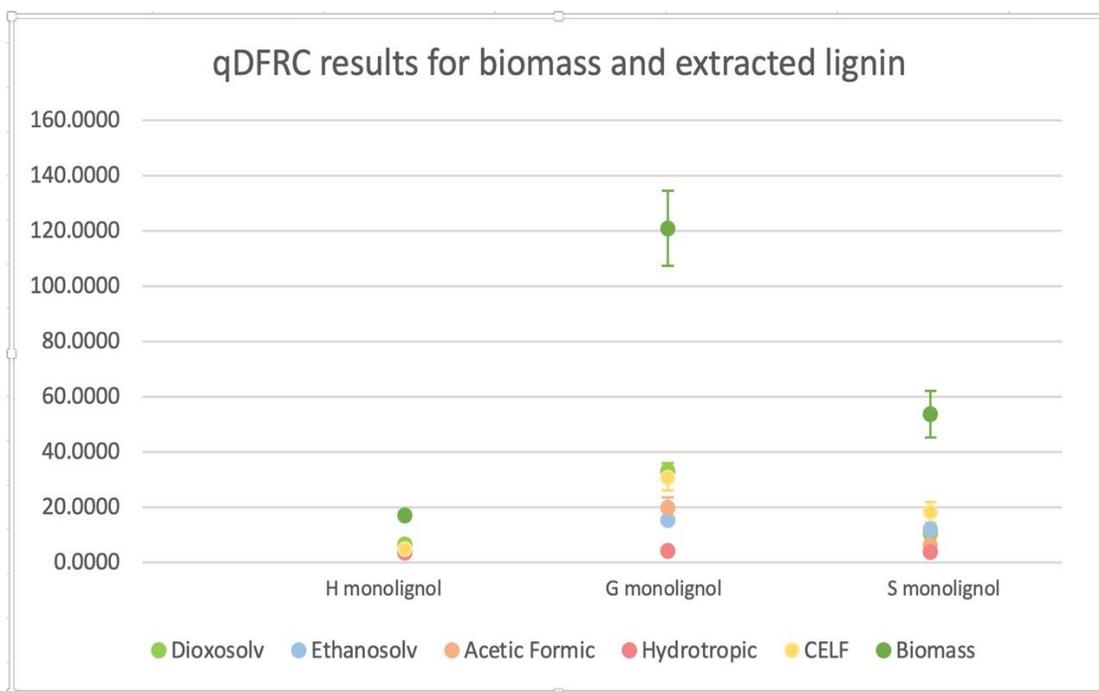
Figure 4.3 Plotted qDFRC results for extracted lignin with pretreatments and untreated biomass on top and the comparison of pretreatments on the bottom. (*) values for untreated biomass are adjusted with the assumption of 12.25%. The values are per mg of lignin in biomass.

The wheat straw biomass does not contain a significant amount of H monolignol, therefore, the amount of recovered H monolignol didn't have a significant contribution on prioritizing pretreatments. There was slightly more H monolignol observed in dioxosolv pretreatment. The other four pretreatment presented almost the same results for H monolignol.

The amount of recovered G monolignol was varying from 4.3 to 33.0 µg/mg between pretreatments with hydrotropic as the lowest yield for G monolignol and dioxosolv as the highest yield. It was noteworthy that dioxosolv had the lowest yield for extracted lignin, while hydrotropic had the highest yield. It was clear that the pretreatment which yield more lignin had more severe condition and imposed more condensation on lignin. Conversely, pretreatment with lower levels of extracted lignin yielded a higher amount of monolignols. CELF pretreatment is the second highest yield in G monolignol (30.7 µg/mg) and very close to dioxosolv. Acetic/formic acid and ethanosolv were also close together with 19.7 and 15.3 µg/mg, respectively followed by hydrotropic as the least amount of recovered G monolignol (4.3 µg/mg).

The highest amount of S monolignol was obtained following CELF pretreatment at 18.2 µg/mg. Dioxosolv which was the highest yield in G monolignol is ranked third after ethanosolv. S monolignol recovered from ethanosolv and dioxosolv found to be close together at 12.1 and 10.5 µg/mg, respectively. Both acetic/formic acid and hydrotropic produced low yields of S monolignol, exhibiting 6.3 and 3.8 µg/mg, respectively.

4.3.7   Statistical Analysis

As mentioned earlier, each pretreatment was replicated three times for the protocol and three times for the analytical technique(qDFRC) resulting in nine data points per pretreatment. In order to combine these three groups, a one-way ANOVA test were taken by GraphPad for each pretreatment to prove that there was no significant difference between these groups for all three H, G and S monolignol.

Among five pretreatments, one of the replicates of acetic/formic acid pretreatment failed to precipitate any lignin, so the sample size for this pretreatment was reduced to six samples which was two group of three data points. For this pretreatment T-test was taken with GraphPad to prove that there was no significant difference between the two groups for three monolignols.

Table 4.5 ANOVA and T-test between sub-groups for each pretreatment for H, G and S monolignol to prove all nine data points can represent one pretreatment method. (*) groups that showed significant difference.

ANOVA and T-test			
Pretreatment	P-value		
	H monolignol	G monolignol	S monolignol
Dioxosolv	0.8682	0.4323	0.1244
Ethanosolv	0.1233	0.0187*	0.0036*
CELF	0.5835	0.0072*	0.0207*
Acetic/formic acid	0.3802	0.5704	0.0501
Hydrotropic	0.7918	0.5809	0.0176*

The ANOVA and T-tests showed that there was no significant difference in all pretreatments for H monolignol. In G monolignol, two of the pretreatments shows significant difference among the three groups. An analysis of multiple comparisons that show each set of data group differences and P-values are presented in Table 4.6. These multiple comparisons are for the ANOVA tests that was showing significant differences between groups in a single pretreatment. The comparison between three groups results in three P-values which usually two of them are below 0.05 or at least closer to 0.05. The common group between these two lower P-values is considered as outlier and deleted from the data set and the remaining six data points have been combined together. In ethanosolv G monolignol ANOVA test, P-value is 0.0187 which shows difference in the three replicates of the pretreatment. The multiple comparisons between groups reveal that comparing replicate 1 vs 3 and 2 vs 3 have the lowest P-value (0.0420 and 0.0857). The difference between the means of the two groups also shows higher differences for group 3 with the other two group. Therefore group 3 data sets were not included in the data points for ethanosolv pretreatment.

The results for CELF pretreatment G monolignol also showed P-value of 0.0072 in the ANOVA test and multiple comparison between groups were investigated. By comparing groups 1 vs 2 and 1 vs 3, the P-value derived to be 0.0238 and 0.0195, which both are below 0.05 and showed they have significant difference. Thus, group 1 was removed from the data set for CELF pretreatment in G monolignol.

The results for S monolignols also had P-values that showed significant differences in groups for the same pretreatment. The P-value for ethanosolv in S monolignol was 0.0036 and multiple comparison showed difference between group 3 with the other two

140

groups and removed from the group. With the same reasoning, group 1 was removed form

CELF, and group 2 was removed from hydrotropic. All evaluations discussed is based on

the mean of the remaining data representing each pretreatment.

Table 4.6 Multiple comparison between groups that have P-values below 0.05 in ANOVA tests.

G monolignol	Adjusted P-value	Mean difference
Ethanosolv1 vs. Ethanosolv2	0.4642	-1.107
Ethanosolv1 vs. Ethanosolv3	0.0542	6.136
Ethanosolv2 vs. Ethanosolv3	0.0091	7.243

G monolignol	Adjusted P-value	Mean difference
CELF1 vs. CELF2	0.0065	-15.70
CELF1 vs. CELF3	0.0223	-20.67
CELF2 vs. CELF3	0.4585	-4.966

S monolignol	Adjusted P-value	Mean difference
Ethanosolv1 vs. Ethanosolv2	0.1041	-1.613
Ethanosolv1 vs. Ethanosolv3	0.0305	3.011
Ethanosolv2 vs. Ethanosolv3	0.0171	4.624

S monolignol	Adjusted P-value	Mean difference
CELF1 vs. CELF2	0.0180	-8.601
CELF1 vs. CELF3	0.0648	-12.16
CELF2 vs. CELF3	0.5969	-3.562

S monolignol	Adjusted P-value	Mean difference
Hydrotropic1 vs. Hydrotropic2	0.1220	-1.537
Hydrotropic1 vs. Hydrotropic3	0.3024	0.9942
Hydrotropic2 vs. Hydrotropic3	0.0409	2.531

After removing the data points discussed above, an ANOVA test was taken for each monolignol between pretreatments to show that there was significant difference among different pretreatments.

The P-value in all three H, G and S monolignol showed significant difference between pretreatments. The ANOVA test on GraphPad showed P-value for all monolignols as <0.0001. The multiple comparison between pretreatments for all three monolignols is presented in Table 4.7.

Multiple comparison for H monolignol exhibits very low P-value for comparison between dioxosolv and other pretreatments. Ethanosolv and CELF pretreatment comparison results in P-value of 0.0449 which mean these two pretreatments have significant difference even though the P-value is close to 0.05. The comparison between ethanosolv vs. acetic/formic acid pretreatment, ethanosolv vs. hydrotropic and CELF vs. acetic/formic acid showed no significant difference. Hydrotropic showed significant difference with both CELF and acetic/formic acid.

The P-value in G monolignol showed significant difference between dioxosolv and other pretreatment except CELF as it can be observed in Figure 4.3 that the error bar of these two pretreatments is overlapping. CELF pretreatment also showed significant difference with all other pretreatments. dioxosolv and CELF pretreatment are the highest yield for G monolignol. Hydrotropic, as the lowest yield of G monolignol also have significant difference with all other pretreatments. Ethanosolv and acetic/formic acid showed no significant difference and had P-value of 0.2204.

Dioxosolv yielded lower S monolignol as is closer to same value for ethanosolv pretreatment. Statistical P-value also showed no significant difference by comparing these

143

two pretreatments. CELF pretreatment had the highest yield for S monolignol and showed significant difference with all other pretreatments, except ethanosolv which showed P-value of 0.0573. Acetic/formic acid and hydrotropic had the lowest yield for S monolignol and showed significant difference with all other pretreatments. These two pretreatments showed no significant difference with each other with a P-value of 0.2506.

Table 4.7 ANOVA test on H, G and S monolignol in pretreatments

H monolignol	Adjusted P-value	Mean difference
Dioxosolv vs. Ethanosolv	<0.0001	2.597
Dioxosolv vs. CELF	0.0002	1.720
Dioxosolv vs. Acetic/formic acid	<0.0001	1.990
Dioxosolv vs. Hydrotropic	<0.0001	3.017
Ethanosolv vs. CELF	0.0449	-0.8775
Ethanosolv vs. Acetic/formic acid	0.0723	-0.6074
Ethanosolv vs. Hydrotropic	0.3986	0.4201
CELF vs. Acetic/formic acid	0.9658	0.2701
CELF vs. Hydrotropic	0.0031	1.298
Acetic/formic acid vs. Hydrotropic	0.0042	1.028

G monolignol	Adjusted P-value	Mean difference
Dioxosolv vs. Ethanosolv	<0.0001	17.63
Dioxosolv vs. CELF	0.9347	2.292
Dioxosolv vs. Acetic/formic acid	0.0005	13.24
Dioxosolv vs. Hydrotropic	<0.0001	28.72
Ethanosolv vs. CELF	0.0014	-15.34
Ethanosolv vs. Acetic/formic acid	0.2204	-4.393
Ethanosolv vs. Hydrotropic	<0.0001	11.09
CELF vs. Acetic/formic acid	0.0095	10.95
CELF vs. Hydrotropic	0.0002	26.43
Acetic/formic acid vs. Hydrotropic	0.0014	15.48

Table 4.7 (Continued)

S monolignol	Adjusted P-value	Mean difference
Dioxosolv vs. Ethanosolv	0.1608	-1.541
Dioxosolv vs. CELF	0.0199	-7.712
Dioxosolv vs. Acetic/formic acid	0.0307	4.212
Dioxosolv vs. Hydrotropic	<0.0001	6.739
Ethanosolv vs. CELF	0.0573	-6.171
Ethanosolv vs. Acetic/formic acid	0.0066	5.752
Ethanosolv vs. Hydrotropic	<0.0001	8.279
CELF vs. Acetic/formic acid	0.0014	11.92
CELF vs. Hydrotropic	0.0017	14.45
Acetic/formic acid vs. Hydrotropic	0.2506	2.527

4.4   Conclusion

The results from qDFRC provided the quantity of liberated monolignol from lignin. This method is selectively cleaving β-ethers. Condensation of lignin occur by forming C-C bonds. Pretreatment methods may result in a breakdown of β-O-4 bonds to some extent and initiate condensation of lignin which reduce this bonding moiety in extracted lignin. It is also possible that forming C-C bonds make the β-O-4 bonds less accessible for qDFRC treatment. It was concluded that the lower recovered monolignols by qDFRC were attributed to lignin condensation.

Lignin in wheat straw biomass was extracted with five different pretreatment methods. Dioxosolv, ethanosolv, CELF, acetic/formic acid and hydrotropic pretreatments were applied to extract the lignin. These pretreatments were evaluated by the efficiency in extracting the lignin with gravimetric analysis and the quality of the extracted lignin using qDFRC as a mean to quantify accessible β-O-4 bonds and the extent of condensation.

All pretreatment showed reduction in the liberated monolignols compared to untreated biomass. Hydrotropic pretreatment showed the highest level of extraction of lignin from biomass however the extracted lignin did not provide high yields for liberated monolignols. The results implied that the hydrotropic pretreatment yields in more condensed lignin even though it can provide more lignin from biomass. Dioxosolv on the other hand provided the least quantity of the extracted lignin from biomass, however the extracted lignin yielded the highest quantity of liberated monolignols.

Based on gravimetric analysis and qDFRC results, it was concluded that CELF pretreatment is the most effective pretreatment in regards of extracting the most quantity of lignin from biomass and retaining the highest content of β-O-4 bond type moiety in the

extracted lignin, compared to other studied pretreatments. As a result of CELF pretreatment, G monolignol had the second highest yield, which is not considerably different from dioxosolv as the highest yield. It also exhibited the highest yield in S monolignol compared to other pretreatments.

In terms of quantity of extracted lignin, with CELF pretreatment, 8.02% lignin extracted from wheat straw biomass which is less than hydrotropic and ethanosolv pretreatment with 12.25% and 11.97%. However, considering the extent of condensation that imposed on the lignin with those pretreatments, CELF proved to be a better pretreatment to extract lignin.

CHAPTER 5. CONCLUSION

Lignin is a complex polymer that is found in plant cell walls and is known to be the most abundant aromatic polymer in nature. Despite its potential value, lignin has traditionally been viewed as a waste product in the pulp and paper industry. Recently, however, lignin has drawn the attention of researchers worldwide who are interested in finding ways to valorize it into possible pharmaceutical and chemical synthons.

Despite its potential, the recalcitrant nature of lignin presents a significant challenge to its valorization. One of the main challenges in the process of lignin valorization is finding an ideal method to break down this valuable polymer into its constituents as low molecular weight aromatic chemicals. Another challenge is the development of ideal analytical tools to characterize the resulting breakdown products. Nuclear Magnetic Resonance (NMR) spectroscopy is extensively used for the analysis of lignin products. Specifically, 2D HSQC provides good information on the bond types and composition of the lignin sample. However, NMR spectroscopy requires a long data acquisition time unless a cryoprobe is used. Additionally, HSQC results are usually complicated and hard to interpret due to the complex nature of lignin. Therefore, there is a critical need to develop a rapid and accurate analytical technique to study the structural information of lignin.

Mass spectrometry is a versatile and powerful tool for the analysis and characterization of a wide range of analytes, including low molecular weight compounds, as well as biomacromolecules such as proteins and peptides. The development of high-end mass analyzers, such as orbitrap, has made it possible to perform high-resolution mass spectrometry (HRMS) for the structural analysis of lignin. Compared to other analytical

techniques, mass spectrometry provides valuable information in a relatively short amount of time. In contrast to NMR spectroscopy, mass spectrometry does not necessarily require a high quantity of the sample mass.

One of the key advantages of mass spectrometry over other analytical techniques is its high sensitivity. Mass spectrometry is capable of detecting trace amounts of analytes in complex mixtures, making it a valuable tool for the analysis of natural products such as lignin. Additionally, mass spectrometry can provide detailed structural information on the analytes of interest, including molecular weight and fragmentation patterns. The combination of high sensitivity and detailed structural analysis makes mass spectrometry a powerful tool for the analysis of lignin and can lead to the development of new strategies for the valorization of this abundant natural resource.

In order to apply mass spectrometry for the structural analysis of lignin, it is essential to use advanced model compounds that can closely resemble the complex structure of natural lignin. These model compounds should contain similar functional groups and linkages as found in natural lignin, enabling accurate identification of the structural characteristics of the lignin sample. However, currently, there is a significant gap in commercially available advanced model compounds. Therefore, the development and use of appropriate model compounds are crucial for the successful application of mass spectrometry in lignin analysis.

The aim of this dissertation is to tackle the challenges associated with lignin analysis by developing advanced model compounds that can accurately mimic the complex structure of natural lignin, and by developing mass spectrometric methods to perform structural analysis of the synthesized model compounds. The ultimate goal of this

work is to provide a better understanding of the structural characteristics of lignin and lignin breakdown products.

Chapter 2 of this dissertation reports the synthesis of three distinct precursors that can be utilized for the production of β-O-4 lignin model dimers or oligomers with a desired sequence. These model compounds have the potential to greatly contribute to the sequential structural analysis of β-O-4 oligomers. The precursors consist of a phenol-terminus, a repeating middle ring that can be used for oligomer synthesis, and a terminating aliphatic-terminus. Each precursor was synthesized for the H, G, and S units, allowing for the synthesis of a varying sequence of oligomers. The precursors were designed to maintain all the functional groups present in natural lignin, including the phenolic group, the hydroxyl groups, and most importantly, the unsaturated alpha-beta carbon bond on the aliphatic chain. The precursors can be coupled with an aldol reaction, followed by reduction and deprotection to produce the desired oligomers. To synthesize a β-O-4 dimer, a phenol-terminus precursor can be coupled with an aliphatic-terminus precursor. Moreover, for synthesis of trimers or oligomers, a phenol-terminus precursor can be coupled with a middle ring precursor followed by deprotection to get the aldehyde group back for the consecutive coupling. The synthesis ends with the coupling of the dimeric unit with an aliphatic-terminus precursor follow by reduction and deprotection for retain the hydroxy group which are the functional groups on natural lignin.

During the course of the synthetic pathway, an intriguing byproduct was discovered and characterized, yielding valuable insights for researchers struggling with similar synthetic scenarios, thereby facilitating the design of model compound synthetic routes.

The synthesized precursors were characterized using gas chromatography-mass spectrometry, with emphasis on the discussion of the EI fragmentation patterns. Additionally, two modified G-βO4-G dimers were synthesized and analyzed using GC-MS. The characterization of these modified dimers, in comparison to their unmodified counterpart, provided significant sequential structural insights. Specifically, it was observed that the fragmentation yielded sequence-specific fragments for ring A and ring B, thus enabling further investigation into the sequence of β-O-4 dimers.

In summary, the synthesis of these model compound precursors and structural analysis with EI fragmentation, will greatly facilitate the sequential structural analysis of lignin dimers and oligomers and contribute to a better understanding of the complex structure of natural lignin. This chapter lays the foundation for the following chapter to synthesize a mix linkage trimer with β-O-4 and β-5 bond types.

Chapter 3 of this dissertation presents the application of the synthetic route developed in the previous chapter for the synthesis of a mix linkage trimer including both β-O-4 and β-5 bond types, starting with a G-β5-G dimer. This trimer, containing the two most abundant bond types in lignin, can offer valuable insights into the structural characterization of lignin. The synthesized trimer was analyzed using a Q-Exactive orbitrap high-resolution mass spectrometer, enabling detailed structural analysis of the lignin oligomer. The lithiated trimer was subjected to HCD tandem mass spectrometry, and the fragmentation patterns were discussed in detail for their structural information. The obtained accurate mass for the fragments disproved three proposed fragment ions, highlighting the importance of high-resolution accurate mass in structural analysis. This chapter not only contributes to the lignin structural investigation using lithium adduct

ionization, which was adapted by our research group but also offers new insights into the synthetic pathways which can be employed to synthesized more advanced model compounds that are essential in lignin characterization research.

According to the results of this investigation, the utilization of lithium adduct ionization in positive ion mode mass spectrometry demonstrated its effectiveness in the characterization of advanced lignin trimer featuring β-O-4 and β-5 bonding motifs. This approach can facilitate the analysis and sequencing of unknown lignin-derived structures via (+) ESI tandem mass spectrometry and thus provide a valuable tool for researchers in the field.

In Chapter 4 of this dissertation, an analytical method of quantitative Derivatization Followed by Reductive Cleavage (qDFRC) was utilized in conjunction with gas chromatograph-mass spectrometry to evaluate the effectiveness of five distinct pretreatment strategies. The objective of this study was to gain insight into the efficacy of these pretreatments in terms of their impact on the lignin composition and structure. The use of qDFRC allowed for the quantitative assessment of the individual lignin monomeric units released during the reductive cleavage step, thus enabling a detailed analysis of the compositional and structural changes in the lignin samples resulting from each pretreatment.

Various sources were hypothesized as contributing factors to the reduction in liberated monolignols observed during the gas chromatography-mass spectrometry characterization employing qDFRC. The first potential source was structural alteration of lignin during pretreatment, which resulted in the loss of alpha hydroxy groups that are supposed to be brominated in the acetyl bromide digestion of lignin and the brominated

alpha carbon is required in the reductive cleavage step to yield monolignol. This alteration could have prevented the liberation of monolignols and consequently reduced the signals observed in the results. The second possible source was the cleavage of some of the β-O-4 bonds during the pretreatment process, which would have resulted in a decrease in the monolignol signals observed. A third potential source was the condensation of lignin during pretreatment, which would have resulted in C-C bond formation and cross-linked complex lignin. While the first two hypotheses could not be rejected, the literature suggests that condensation is commonly observed in extracted lignin. Therefore, the third hypothesis was considered as a potential cause for the reduction in liberated monolignols.

The characterization of different pretreatment methods can be important in the field of biomass conversion, where lignocellulosic biomass is converted into biofuels, bioproducts, and other value-added chemicals. Pretreatment is an essential step in this process, which involves modifying the structure of the biomass to make it more accessible for further conversion. The use of qDFRC with GC-MS characterization can provide valuable information on the structural changes that occur during pretreatment and can help researchers optimize the process for maximum efficiency and product yields. Therefore, this chapter may have implications for the development of more sustainable and efficient bioconversion processes.

This dissertation has presented significant insights into the structural analysis of lignin using advanced mass spectrometric and synthetic techniques. While significant progress has been made, there are still areas for future research to further enhance our understanding of the complex structure of lignin. The synthesis and fragmentation study of β-O-4 oligomers using the synthesized precursors can contribute to the structural

elucidation of lignin breakdown products. Additionally, the use of isotope-labeled model compounds with β-5 bond type can provide further understanding of the fragmentation behavior of this bond type, which is not well-studied in the literature. The synthesis of more advanced model compounds with mixed linkage using the synthetic pathway presented can also provide additional insights into the structure of lignin. Moreover, the use of qDFRC to optimize CELF pretreatment conditions, coupled with other analytical techniques, can improve the quality of extracted lignin with this pretreatment method. These future research directions can contribute to the ultimate goal of valorization of lignin as a renewable and sustainable resource for various industrial applications.

In conclusion, this dissertation has significantly contributed to the development of advanced lignin model compounds synthesis and characterization using mass spectrometric techniques for the structural analysis of lignin, with the long-term aim of facilitating the valorization of lignin as a renewable and cost-effective resource for the production of high-value compounds. By employing multifaceted techniques such as high-resolution mass spectrometry and quantitative derivatization followed by reductive cleavage, valuable insights into the chemical structure of lignin and its degradation products have been gained. The findings provide a foundation for future studies aimed at characterization of unknown lignin samples and improving lignin extraction and processing techniques. The future research directions presented in this dissertation have the potential to advance the production of value-added products from lignin, which is essential for building a more sustainable and prosperous future. Through the exploration of the structural complexities of lignin, this research aims to contribute to the development

of sustainable technologies for the production of biochemicals and biomaterials, and ultimately create a more sustainable future for generations to come.

APPENDIX 1. EI mass spectrum of G unit phenol-terminus precursor, $[M]^{•+}=308$
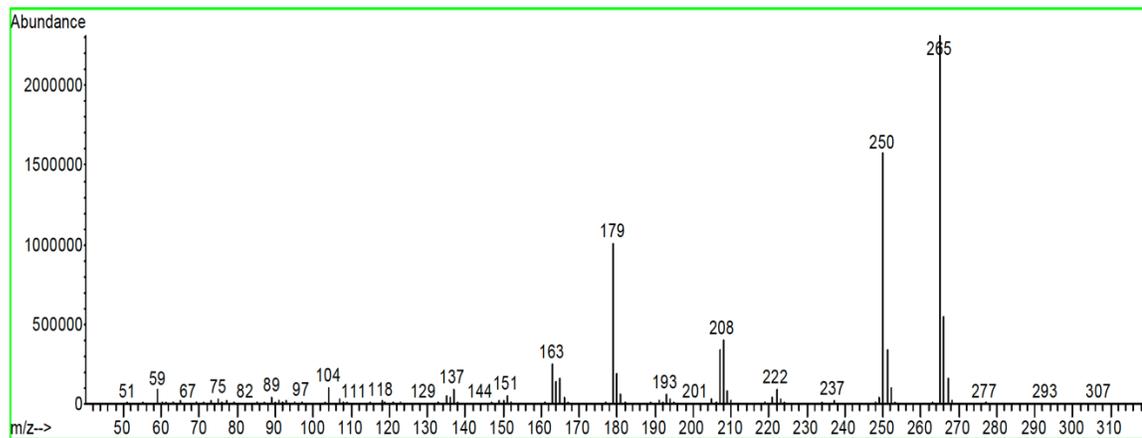
APPENDIX 2. EI mass spectrum of ethyl ferulate [M]$^{\bullet+}$= 294
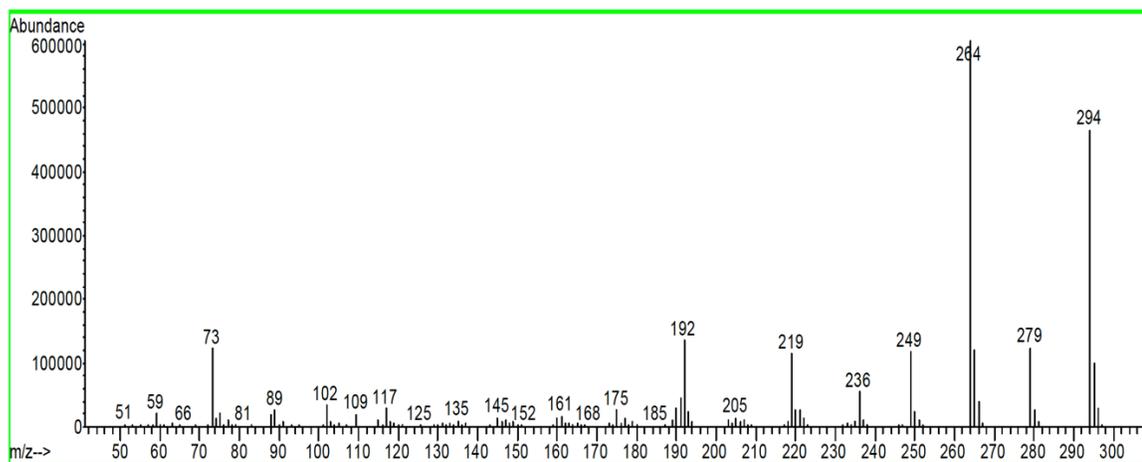
APPENDIX 3. EI mass spectrum of ethyl sinapate $[M]^{\bullet+} = 324$
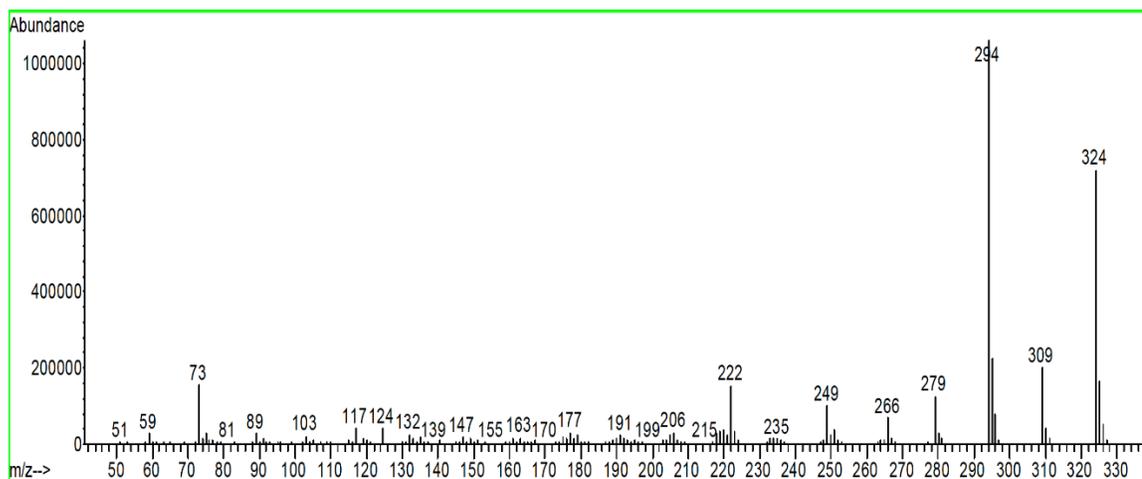
APPENDIX 4. EI mass spectrum of G unit aliphatic-terminus precursor $[M]^{\bullet+}= 308$
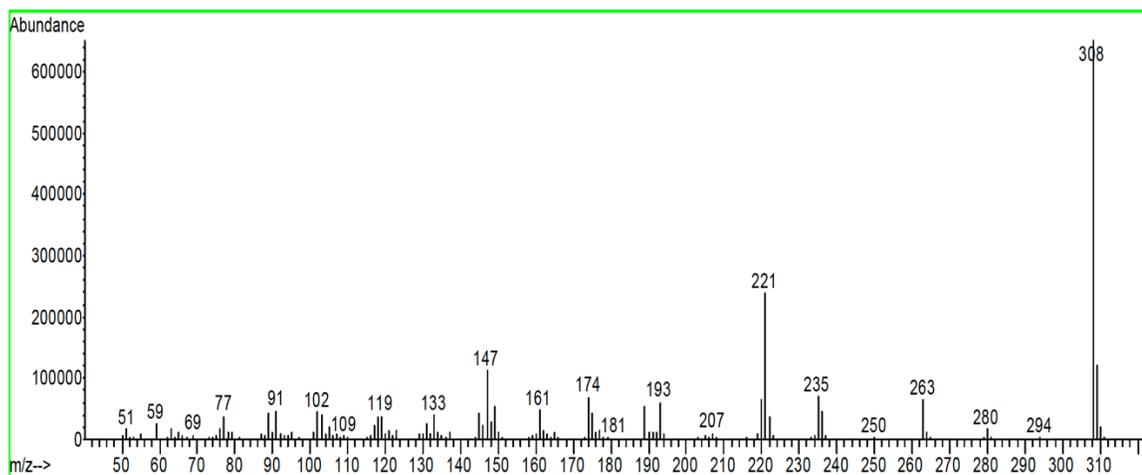
APPENDIX 5. EI mass spectrum of S unit aliphatic-terminus precursor [M]$^{•+}$= 338
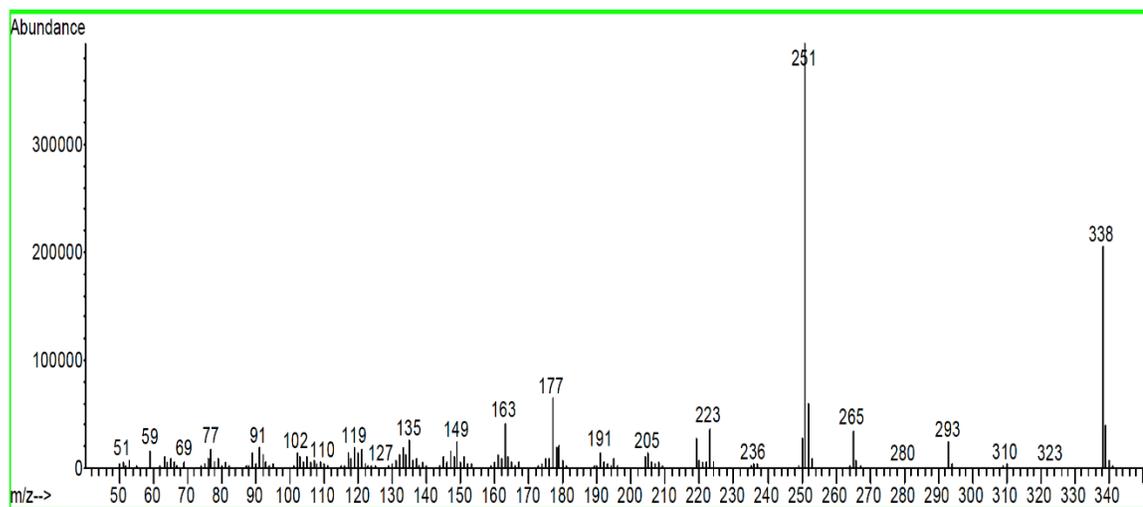
APPENDIX 6. EI mass spectrum of esterified 4-hydroxy benzaldehyde [M]$^{\bullet+}$= 208
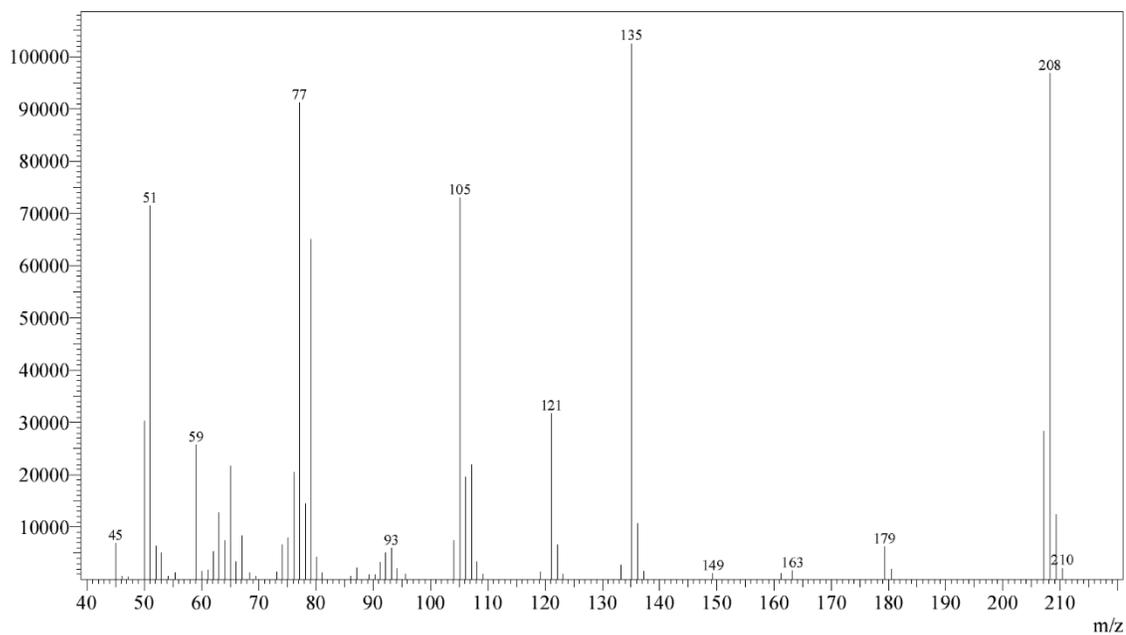
APPENDIX 7. EI mass spectrum of esterified syringe aldehyde $[M]^{\bullet+} = 268$
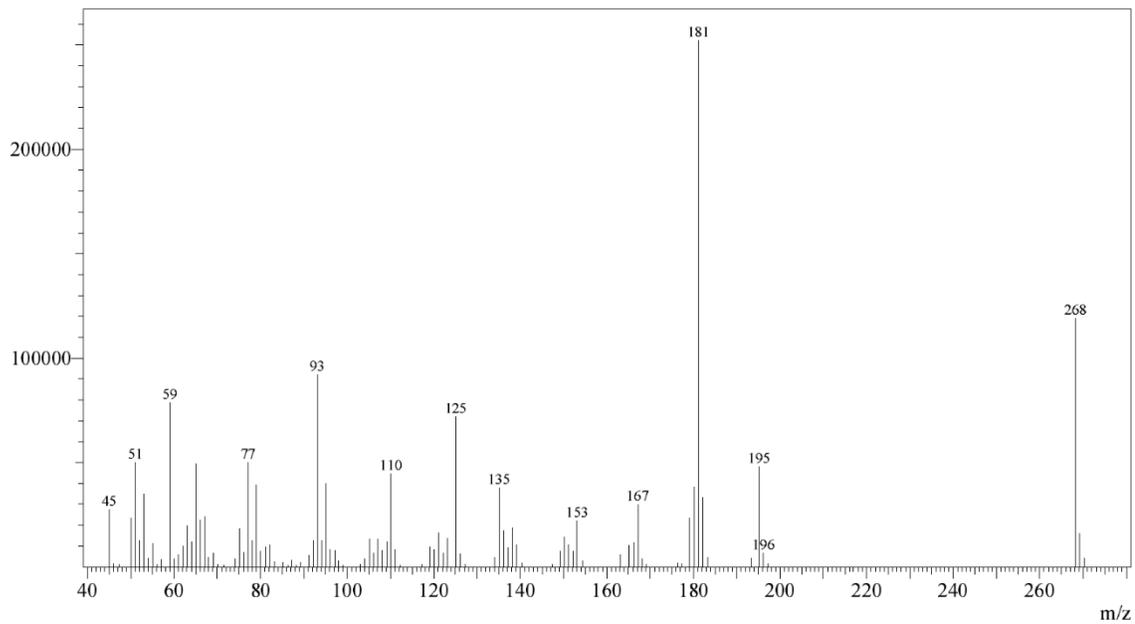
APPENDIX 8. EI mass spectrum of H unit middle ring precursor $[M]^{\bullet+} = 252$
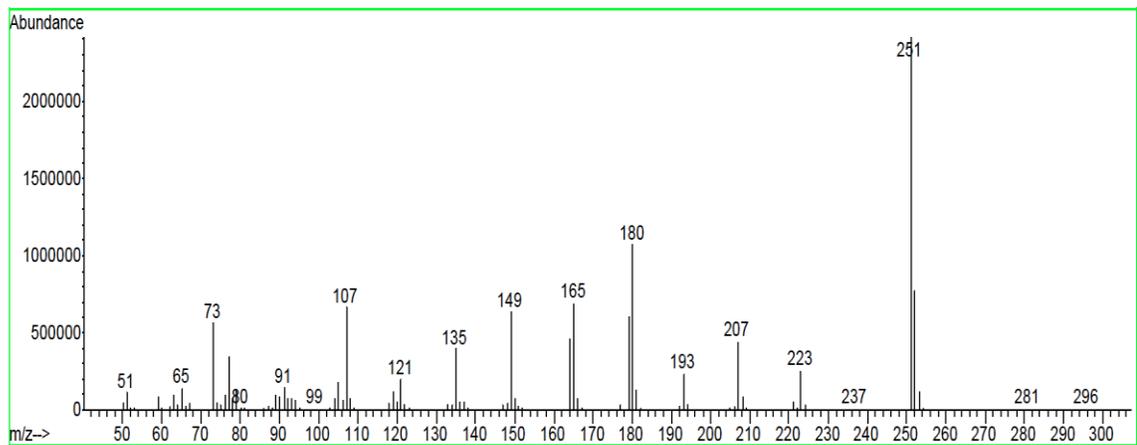
APPENDIX 9. EI mass spectrum of S unit middle ring precursor [M]$^{\bullet+}$= 312
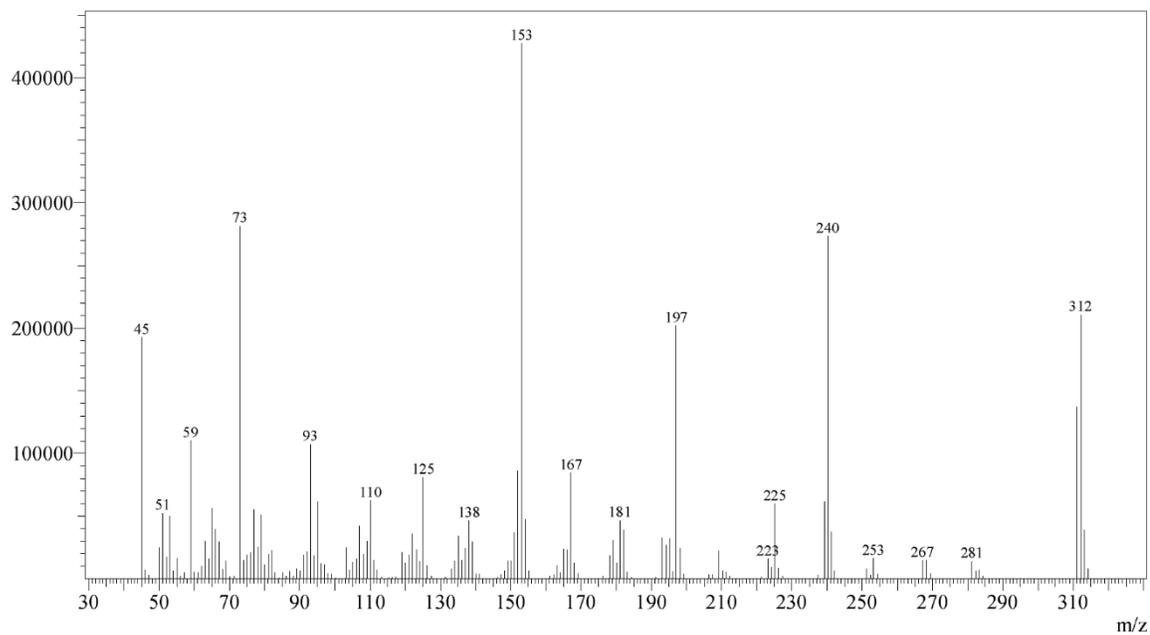
APPENDIX 10. EI mass spectrum of acylated vanillin [M]$^{\bullet+}$= 194
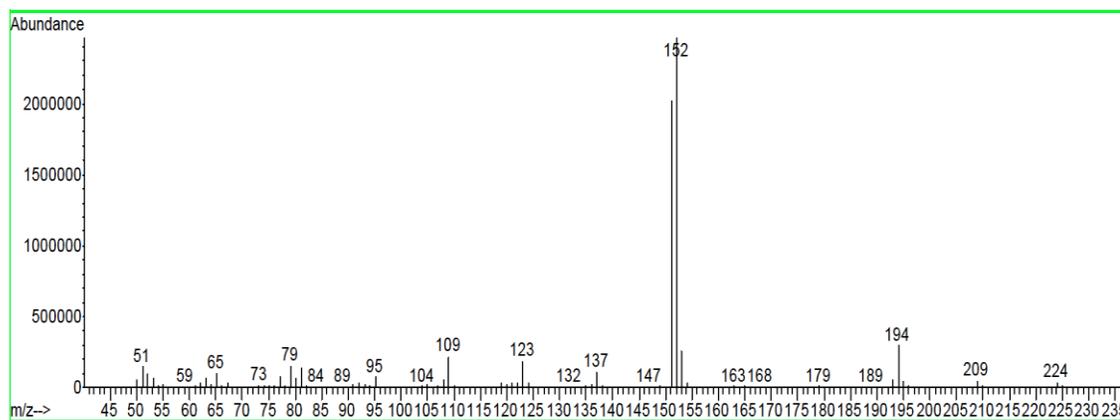
APPENDIX 11. EI mass spectrum of G-β5-G dimer [M]$^{\bullet+}$= 574
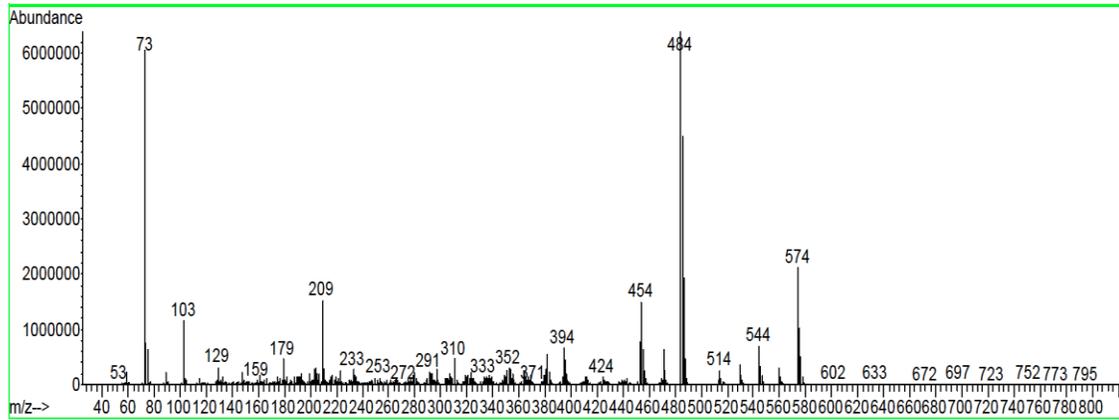
APPENDIX 12. EI mass spectrum of G-β5-G ethyl acetate (compound 7 in Chapter 3) [M]$^{\bullet+}$= 588
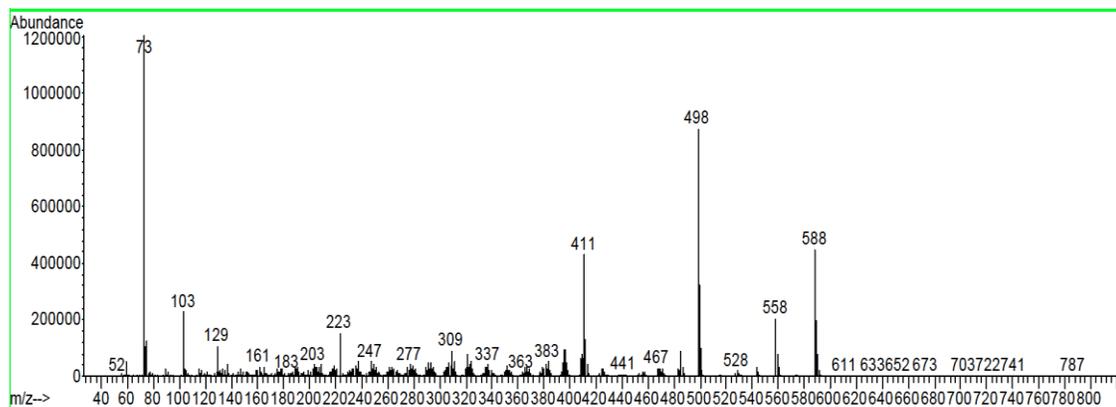
APPENDIX 13. LTQ ESI mass spectrum of compound 8 in Chapter 3 $[M+Li]^+= 535$

B5_EA_Ac_PURE_12_27_21 #162 RT: 1.57 AV: 1 NL: 3.63E6
T: ITMS + p ESI Full ms [145.00-1000.00]
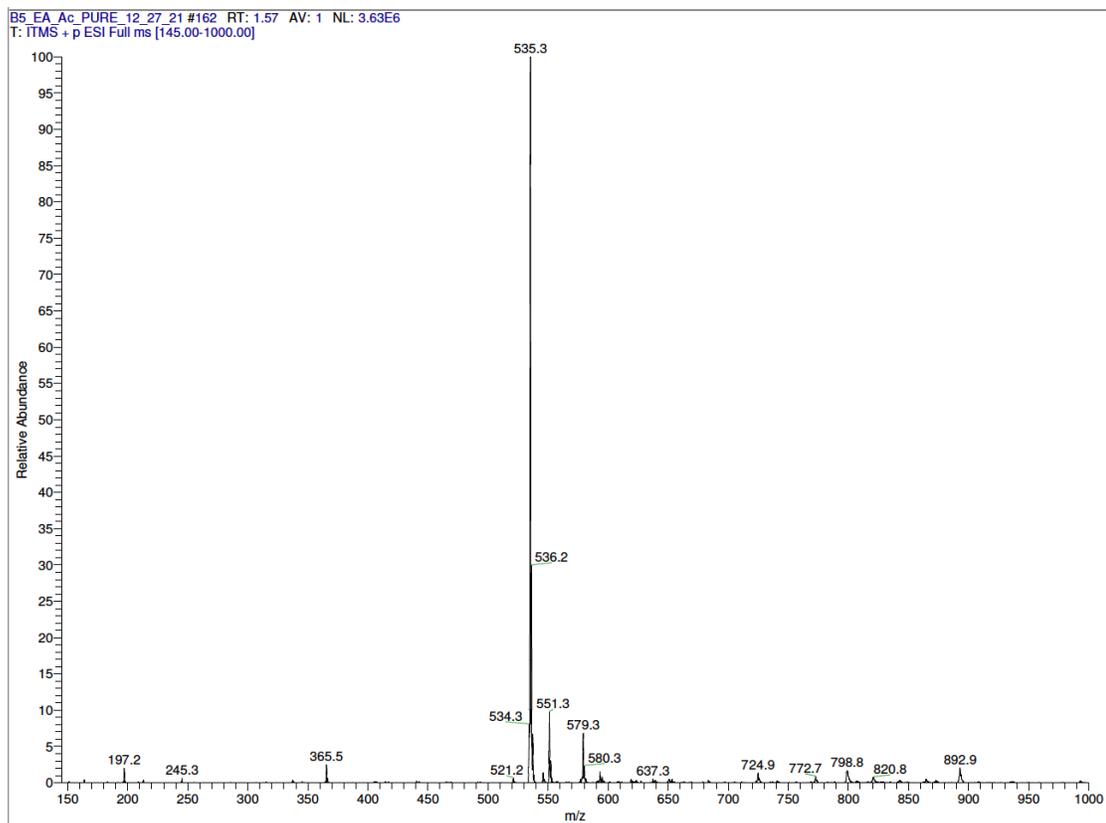
APPENDIX 14. LTQ ESI mass spectrum of compound 9 in Chapter 3 [M+Li]$^+$= 729
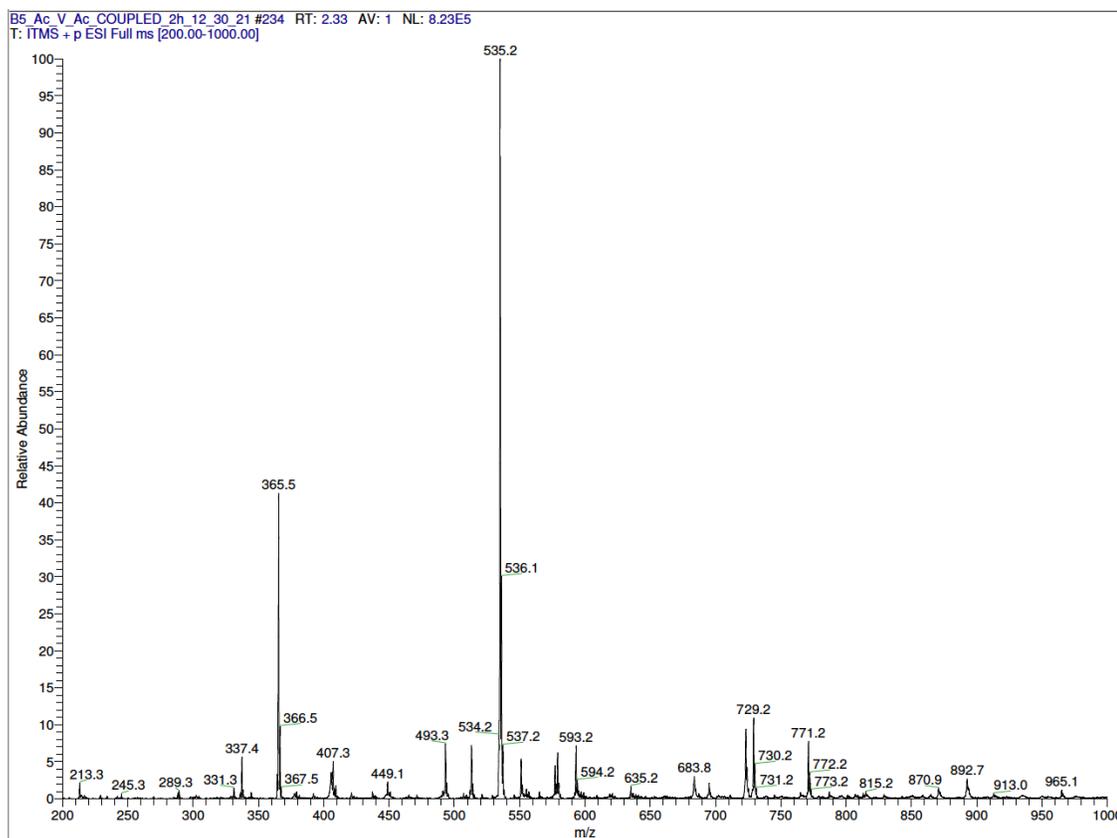
REFERENCES

Uncategorized References

(1) Asare, S. O.; Kamali, P.; Huang, F.; Lynn, B. C. Application of chloride adduct ionization tandem mass spectrometry for characterizing and sequencing synthetic lignin model compounds. *Energy & fuels* **2018**, *32* (5), 5990-5998.

(2) Xu, C.; Arancon, R. A. D.; Labidi, J.; Luque, R. Lignin depolymerisation strategies: towards valuable chemicals and fuels. *Chemical Society Reviews* **2014**, *43* (22), 7485-7500.

(3) Chandel, A. K.; Garlapati, V. K.; Singh, A. K.; Antunes, F. A. F.; da Silva, S. S. The path forward for lignocellulose biorefineries: bottlenecks, solutions, and perspective on commercialization. *Bioresource Technology* **2018**, *264*, 370-381.

(4) Ragauskas, A. J.; Williams, C. K.; Davison, B. H.; Britovsek, G.; Cairney, J.; Eckert, C. A.; Frederick Jr, W. J.; Hallett, J. P.; Leak, D. J.; Liotta, C. L. The path forward for biofuels and biomaterials. *science* **2006**, *311* (5760), 484-489.

(5) Jung, S.-J.; Kim, S.-H.; Chung, I.-M. Comparison of lignin, cellulose, and hemicellulose contents for biofuels utilization among 4 types of lignocellulosic crops. *Biomass and bioenergy* **2015**, *83*, 322-327.

(6) Sun, X.-B.; Gao, D.-Y.; Cao, J.-W.; Liu, Y.; Rong, Z.-T.; Wang, J.-K.; Wang, Q. BsLPMO10A from Bacillus subtilis boosts the depolymerization of diverse polysaccharides linked via β-1, 4-glycosidic bonds. *International Journal of Biological Macromolecules* **2023**, 123133.

(7) Kumar, N.; Dixit, A. Chapter 4-Management of biomass. Elsevier: 2021; pp 97-140.

(8) Ruiz-Dueñas, F. J.; Martínez, Á. T. Microbial degradation of lignin: how a bulky recalcitrant polymer is efficiently recycled in nature and how we can take advantage of this. *Microbial biotechnology* **2009**, *2* (2), 164-177.

(9) Ciofi-Baffoni, S.; Banci, L.; Brandi, A. Synthesis of oligomeric mimics of lignin. *Journal of the Chemical Society, Perkin Transactions 1* **1998**, (19), 3207-3218.

(10) Wu, W.; Dutta, T.; Varman, A. M.; Eudes, A.; Manalansan, B.; Loqué, D.; Singh, S. Lignin valorization: two hybrid biochemical routes for the conversion of polymeric lignin into value-added chemicals. *Scientific reports* **2017**, *7* (1), 8420.

(11) U.S. energy facts explained - Consumption and production. (EIA), U. S. E. I. A., Ed.

(12) Chio, C.; Sain, M.; Qin, W. Lignin utilization: A review of lignin depolymerization from various aspects. *Renewable and Sustainable Energy Reviews* **2019**, *107*, 232-249.

(13) Ten, E.; Vermerris, W. Recent developments in polymers derived from industrial lignin. *Journal of Applied Polymer Science* **2015**, *132* (24).

(14) Boerjan, W.; Ralph, J.; Baucher, M. Lignin biosynthesis. *Annual review of plant biology* **2003**, *54* (1), 519-546.

(15) Labeeuw, L.; Martone, P. T.; Boucher, Y.; Case, R. J. Ancient origin of the biosynthesis of lignin precursors. *Biology Direct* **2015**, *10* (1), 1-21.

(16) Liu, X.; Bouxin, F. P.; Fan, J.; Budarin, V. L.; Hu, C.; Clark, J. H. Recent advances in the catalytic depolymerization of lignin towards phenolic chemicals: a review. *ChemSusChem* **2020**, *13* (17), 4296-4317.

(17) Schutyser, W.; Renders, a. T.; Van den Bosch, S.; Koelewijn, S.-F.; Beckham, G.; Sels, B. F. Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chemical Society Reviews* **2018**, *47* (3), 852-908.

(18) Li, C.; Zhao, X.; Wang, A.; Huber, G. W.; Zhang, T. Catalytic transformation of lignin for the production of chemicals and fuels. *Chemical reviews* **2015**, *115* (21), 11559-11624.

(19) Bruijnincx, P. C.; Weckhuysen, B. M. Shale gas revolution: an opportunity for the production of biobased chemicals? *Angewandte Chemie International Edition* **2013**, *52* (46), 11980-11987.

(20) Zakzeski, J.; Bruijnincx, P. C.; Jongerius, A. L.; Weckhuysen, B. M. The catalytic valorization of lignin for the production of renewable chemicals. *Chemical reviews* **2010**, *110* (6), 3552-3599.

(21) Capanema, E. A.; Balakshin, M. Y.; Kadla, J. F. A comprehensive approach for quantitative lignin characterization by NMR spectroscopy. *Journal of agricultural and food chemistry* **2004**, *52* (7), 1850-1860.

(22) Wen, J.-L.; Sun, S.-L.; Xue, B.-L.; Sun, R.-C. Recent advances in characterization of lignin polymer by solution-state nuclear magnetic resonance (NMR) methodology. *Materials* **2013**, *6* (1), 359-391.

(23) Shi, Z.; Xu, G.; Deng, J.; Dong, M.; Murugadoss, V.; Liu, C.; Shao, Q.; Wu, S.; Guo, Z. Structural characterization of lignin from D. sinicus by FTIR and NMR techniques. *Green Chemistry Letters and Reviews* **2019**, *12* (3), 235-243.

(24) 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.

(25) Li, S.; Luo, Z.; Wang, W.; Lu, K.; Yang, Y.; Liang, X. Characterization of pyrolytic lignin and insight into its formation mechanisms using novel techniques and DFT method. *Fuel* **2020**, *262*, 116516.

(26) Bayerbach, R.; Nguyen, V. D.; Schurr, U.; Meier, D. Characterization of the water-insoluble fraction from fast pyrolysis liquids (pyrolytic lignin): Part III. Molar mass characteristics by SEC, MALDI-TOF-MS, LDI-TOF-MS, and Py-FIMS. *Journal of analytical and applied pyrolysis* **2006**, *77* (2), 95-101.

(27) Saito, T.; Perkins, J. H.; Vautard, F.; Meyer, H. M.; Messman, J. M.; Tolnai, B.; Naskar, A. K. Methanol fractionation of softwood kraft lignin: Impact on the lignin properties. *ChemSusChem* **2014**, *7* (1), 221-228.

(28) Srivastava, R. K.; Shetti, N. P.; Reddy, K. R.; Kwon, E. E.; Nadagouda, M. N.; Aminabhavi, T. M. Biomass utilization and production of biofuels from carbon neutral materials. *Environmental Pollution* **2021**, *276*, 116731.

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

(30) Modenbach, A. A.; Nokes, S. E. Effects of sodium hydroxide pretreatment on structural components of biomass. *Transactions of the ASABE* **2014**, *57* (4), 1187-1198.

(31) Xu, J.; Liu, B.; Wu, L.; Hu, J.; Hou, H.; Yang, J. A waste-minimized biorefinery scenario for the hierarchical conversion of agricultural straw into prebiotic xylooligosaccharides, fermentable sugars and lithium-sulfur batteries. *Industrial Crops and Products* **2019**, *129*, 269-280.

(32) 4.0, C. B. Pulp and Paper. 2022.

(33) Sun, C.; Song, G.; Pan, Z.; Tu, M.; Kharaziha, M.; Zhang, X.; Show, P.-L.; Sun, F. Advances in organosolv modified components occurring during the organosolv pretreatment of lignocellulosic biomass. *Bioresource Technology* **2022**, 128356.

(34) Achinivu, E. C.; Mohan, M.; Choudhary, H.; Das, L.; Huang, K.; Magurudeniya, H. D.; Pidatala, V. R.; George, A.; Simmons, B. A.; Gladden, J. M. A predictive toolset for the identification of effective lignocellulosic pretreatment solvents: a case study of solvents tailored for lignin extraction. *Green Chemistry* **2021**, *23* (18), 7269-7289.

(35) Huang, D.; Li, R.; Xu, P.; Li, T.; Deng, R.; Chen, S.; Zhang, Q. The cornerstone of realizing lignin value-addition: Exploiting the native structure and properties of lignin by extraction methods. *Chemical Engineering Journal* **2020**, *402*, 126237.

(36) Madadi, M.; Song, G.; Karimi, K.; Zhu, D.; Elsayed, M.; Sun, F.; Abomohra, A. One-step lignocellulose fractionation using acid/pentanol pretreatment for enhanced fermentable sugar and reactive lignin production with efficient pentanol retrievability. *Bioresource Technology* **2022**, *359*, 127503.

(37) Questell-Santiago, Y. M.; Zambrano-Varela, R.; Talebi Amiri, M.; Luterbacher, J. S. Carbohydrate stabilization extends the kinetic limits of chemical polysaccharide depolymerization. *Nature chemistry* **2018**, *10* (12), 1222-1228.

(38) Devendra, L. P.; Kumar, M. K.; Pandey, A. Evaluation of hydrotropic pretreatment on lignocellulosic biomass. *Bioresource Technology* **2016**, *213*, 350-358.

(39) Yin, C.; Wang, M.; Ma, Q.; Bian, H.; Ren, H.; Dai, H.; Cheng, J. Valorization of Rice Straw via Hydrotropic Lignin Extraction and Its Characterization. *Molecules* **2021**, *26* (14), 4123.

(40) Abdelkafi, F.; Ammar, H.; Rousseau, B.; Tessier, M.; El Gharbi, R.; Fradet, A. Structural analysis of alfa grass (Stipa tenacissima L.) lignin obtained by acetic acid/formic acid delignification. *Biomacromolecules* **2011**, *12* (11), 3895-3902.

(41) Wang, T.-P.; Li, H.; Yuan, J.-M.; Li, W.-X.; Li, K.; Huang, Y.-B.; Xiao, L.-P.; Lu, Q. Structures and pyrolytic characteristics of organosolv lignins from typical softwood, hardwood and herbaceous biomass. *Industrial Crops and Products* **2021**, *171*, 113912.

(42) Bowman, A. S.; Asare, S. O.; Lynn, B. C. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis for characterization of lignin oligomers using cationization techniques and 2, 5-dihydroxyacetophenone (DHAP) matrix. *Rapid Communications in Mass Spectrometry* **2019**, *33* (8), 811-819.

(43) Kumar, P.; Barrett, D. M.; Delwiche, M. J.; Stroeve, P. Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Industrial & engineering chemistry research* **2009**, *48* (8), 3713-3729.

(44) Finehout, E. J.; Lee, K. H. An introduction to mass spectrometry applications in biological research. *Biochemistry and molecular biology Education* **2004**, *32* (2), 93-100.

(45) Burlingame, A. L.; Carr, S. A.; Baldwin, M. A. *Mass spectrometry in biology & medicine*; Springer Science & Business Media, 1999.

(46) Han, X.; Aslanian, A.; Yates III, J. R. Mass spectrometry for proteomics. *Current opinion in chemical biology* **2008**, *12* (5), 483-490.

(47) Wang, Y.; Sun, J.; Qiao, J.; Ouyang, J.; Na, N. A "soft" and "hard" ionization method for comprehensive studies of molecules. *Analytical chemistry* **2018**, *90* (24), 14095-14099.

(48) Maciel, E. V. S.; Pereira dos Santos, N. G.; Vargas Medina, D. A.; Lanças, F. M. Electron ionization mass spectrometry: Quo vadis? *Electrophoresis* **2022**, *43* (15), 1587-1600.

(49) Poole, C. *Gas chromatography*; Elsevier, 2021.

(50) Kebarle, P.; Verkerk, U. H. On the mechanism of electrospray ionization mass spectrometry (ESIMS). *Electrospray and MALDI mass spectrometry: fundamentals, instrumentation, practicalities, and biological applications* **2010**, 1-48.

(51) Page, J. S.; Kelly, R. T.; Tang, K.; Smith, R. D. Ionization and transmission efficiency in an electrospray ionization—mass spectrometry interface. *Journal of the American Society for Mass Spectrometry* **2007**, *18* (9), 1582-1590.

(52) Asare, S. O.; Lynn, B. C. A comparative study of the electrospray ionization response of β-O-4′ lignin model compounds. *Journal of Mass Spectrometry* **2019**, *54* (6), 540-548.

(53) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the mechanism of electrospray ionization. ACS Publications: 2013.

(54) Banerjee, S.; Mazumdar, S. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *International journal of analytical chemistry* **2012**, *2012*.

(55) Dawson, P. H. *Quadrupole mass spectrometry and its applications*; Elsevier, 2013.

(56) Buckingham, A.; Disch, R.-L.; Dunmur, D. Quadrupole moments of some simple molecules. *Journal of the American Chemical Society* **1968**, *90* (12), 3104-3107.

(57) Zubarev, R. A.; Makarov, A. Orbitrap mass spectrometry. ACS Publications: 2013.

(58) Savaryn, J. P.; Toby, T. K.; Kelleher, N. L. A researcher's guide to mass spectrometry-based proteomics. *Proteomics* **2016**, *16* (18), 2435-2443.

(59) Eliuk, S.; Makarov, A. Evolution of orbitrap mass spectrometry instrumentation. *Annual review of analytical chemistry* **2015**, *8*, 61-80.

(60) Hecht, E. S.; Scigelova, M.; Eliuk, S.; Makarov, A. Fundamentals and advances of orbitrap mass spectrometry. *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation* **2006**, 1-40.

(61) Letourneau, D. R.; Volmer, D. A. Mass spectrometry-based methods for the advanced characterization and structural analysis of lignin: A review. *Mass Spectrometry Reviews* **2023**, *42* (1), 144-188.

(62) Asare, S. O.; Huang, F.; Lynn, B. C. Characterization and sequencing of lithium cationized β-O-4 lignin oligomers using higher-energy collisional dissociation mass spectrometry. *Analytica Chimica Acta* **2019**, *1047*, 104-114.

(63) Njiojob, C. N.; Bozell, J. J.; Long, B. K.; Elder, T.; Key, R. E.; Hartwig, W. T. Enantioselective Syntheses of Lignin Models: An Efficient Synthesis of β-O-4 Dimers and Trimers by Using the Evans Chiral Auxiliary. *Chemistry–A European Journal* **2016**, *22* (35), 12506-12517.

(64) Katahira, R.; Kamitakahara, H.; Takano, T.; Nakatsubo, F. Synthesis of β-O-4 type oligomeric lignin model compound by the nucleophilic addition of carbanion to the aldehyde group. *Journal of wood science* **2006**, *52* (3), 255-260.

(65) Flourat, A. L.; Peru, A. A.; Haudrechy, A.; Renault, J.-H.; Allais, F. First total synthesis of (β-5)-(β-O-4) dihydroxytrimer and dihydrotrimer of coniferyl alcohol (G): advanced lignin model compounds. *Frontiers in Chemistry* **2019**, *7*, 842.

(66) Binder, J. B.; Gray, M. J.; White, J. F.; Zhang, Z. C.; Holladay, J. E. Reactions of lignin model compounds in ionic liquids. *Biomass and Bioenergy* **2009**, *33* (9), 1122-1130.

(67) Yang, L.; Seshan, K.; Li, Y. A review on thermal chemical reactions of lignin model compounds. *Catalysis Today* **2017**, *298*, 276-297.

(68) Huang, F. *Application of high-resolution accurate mass (HRAM) mass spectrometry for analysis of lignin model compounds and the post-pretreatment products*; University of Kentucky, 2017. Theses and Dissertations--Chemistry. 74.
https://uknowledge.uky.edu/chemistry_etds/74

(69) Asare, S. O. Bottom-up Lignomics: Towards the Development of Adduct Electrospray Ionization Mass Spectrometric Methods to Characterize and Sequence Lignin Oligomers. **2019**. Theses and Dissertations--Chemistry. 115.
https://uknowledge.uky.edu/chemistry_etds/115

(70) Moradipour, M.; Tong, X.; Novak, B.; Kamali, P.; Asare, S. O.; Lynn, B. C.; Moldovan, D.; Rankin, S. E.; Knutson, B. L. Interaction of lignin dimers with model cell membranes: A quartz crystal microbalance and molecular dynamics simulation study. *Biointerphases* **2021**, *16* (4), 041003.

(71) Dean, K. R.; Lynn, B. C. Lithium cation basicity estimates of lignin β-O-4 dimers by the kinetic method utilizing a novel ladder approach. *International Journal of Mass Spectrometry* **2020**, *457*, 116416.

(72) Jardine, I.; Fenselau, C. The high pressure nitric oxide mass spectra of aldehydes. *Organic Mass Spectrometry* **1975**, *10* (9), 748-751.

(73) Moreau, R. A.; Sharma, M. E.; Nuñez, A.; Mullen, C. A.; Powell, M. J.; Jones, K.; Harron, A.; Cafmeyer, J. T. Identification of Unique Aldehyde Dimers in Sorghum Wax Recovered after Fermentation in a Commercial Fuel Ethanol Plant. *Journal of the American Oil Chemists' Society* **2020**, *97* (12), 1299-1308.

(74) Lyakhovetsky, Y. I.; Pleshkova, A. P.; Shilova, E. A.; Ponomareva, T. V.; Gasanov, R. G.; Tumanskii, B. L.; Borisov, Y. A.; Nekrasov, Y. S. A comparative study of homolytic reactions of fullerenes with aldehydes in a mass spectrometer under electron impact and in solution under UV irradiation. *European Journal of Mass Spectrometry* **2012**, *18* (4), 361-376.

(75) Abiedalla, Y.; Smith, L. W.; Abdel-Hay, K. M.; Neel, L.; Belal, T. S.; Thaxton-Weissenfluh, A.; Smith, F.; DeRuiter, J.; Clark, C. R. Spectroscopic differentiation and chromatographic separation of regioisomeric indole aldehydes: synthetic cannabinoids precursors. *Forensic Chemistry* **2019**, *12*, 78-90.

(76) Ameta, U.; Ojha, S.; Bhambi, D.; Talesara, G. L. Synthetic studies on some 3-[(5-arylidene-4-oxo-1, 3-thiazolidin-2-yliden) amino]-2-phenylquinazolin-4 (3H)-ones and their ethoxyphthalimide derivatives. *Arkivoc* **2006**, *13*, 83-89.

(77) Heras-Mozos, R.; Hernández, R.; Gavara, R.; Hernández-Muñoz, P. Dynamic covalent chemistry of imines for the development of stimuli-responsive chitosan films as carriers of sustainable antifungal volatiles. *Food Hydrocolloids* **2022**, *125*, 107326.

(78) Liu, L.; Frohn, M.; Xi, N.; Dominguez, C.; Hungate, R.; Reider, P. J. A soluble base for the copper-catalyzed imidazole N-arylations with aryl halides. *The Journal of Organic Chemistry* **2005**, *70* (24), 10135-10138.

(79) *Mass Spectrometry: Fragmentation*.
http://chemistry.miamioh.edu/gung/CHM526/pdfs/Mass-fragmentation.pdf (accessed 2023 04/13/2023).

(80) Zhang, X.; Pugh, J. K.; Ross, P. N. Computation of thermodynamic oxidation potentials of organic solvents using density functional theory. *Journal of The Electrochemical Society* **2001**, *148* (5), E183.

(81) Ceter, N. M. S. D. *Nist Chemistry Webbook*. 2014.
https://webbook.nist.gov/cgi/cbook.cgi?ID=C646060&Mask=200#Mass-Spec (accessed 2023 04/07/2023).

(82) Tong, X.; Moradipour, M.; Novak, B.; Kamali, P.; Asare, S. O.; Knutson, B. L.; Rankin, S. E.; Lynn, B. C.; Moldovan, D. Experimental and molecular dynamics simulation study of the effects of lignin dimers on the gel-to-fluid phase transition in DPPC bilayers. *The Journal of Physical Chemistry B* **2019**, *123* (39), 8247-8260.

(83) Liu, Y.; D'Agostino, L. A.; Qu, G.; Jiang, G.; Martin, J. W. High-resolution mass spectrometry (HRMS) methods for nontarget discovery and characterization of poly-and per-fluoroalkyl substances (PFASs) in environmental and human samples. *TrAC Trends in Analytical Chemistry* **2019**, *121*, 115420.

(84) Lucci, P.; Saurina, J.; Núñez, O. Trends in LC-MS and LC-HRMS analysis and characterization of polyphenols in food. *TrAC Trends in Analytical Chemistry* **2017**, *88*, 1-24.

(85) Han, X.; MacKinnon, M. D.; Martin, J. W. Estimating the in situ biodegradation of naphthenic acids in oil sands process waters by HPLC/HRMS. *Chemosphere* **2009**, *76* (1), 63-70.

(86) Ramanathan, R.; Jemal, M.; Ramagiri, S.; Xia, Y. Q.; Humpreys, W. G.; Olah, T.; Korfmacher, W. A. It is time for a paradigm shift in drug discovery bioanalysis: from SRM to HRMS. *Journal of mass spectrometry* **2011**, *46* (6), 595-601.

(87) Gago-Ferrero, P.; Schymanski, E. L.; Bletsou, A. A.; Aalizadeh, R.; Hollender, J.; Thomaidis, N. S. Extended suspect and non-target strategies to characterize emerging polar organic contaminants in raw wastewater with LC-HRMS/MS. *Environmental science & technology* **2015**, *49* (20), 12333-12341.

(88) Sheng, H.; Tang, W.; Gao, J.; Riedeman, J. S.; Li, G.; Jarrell, T. M.; Hurt, M. R.; Yang, L.; Murria, P.; Ma, X. (−) ESI/CAD MS n Procedure for Sequencing Lignin Oligomers Based on a Study of Synthetic Model Compounds with β-O-4 and 5-5 Linkages. *Analytical chemistry* **2017**, *89* (24), 13089-13096.

(89) Zhang, J.; Jiang, Y.; Easterling, L. F.; Anstner, A.; Li, W.; Alzarieni, K. Z.; Dong, X.; Bozell, J.; Kenttämaa, H. I. Compositional analysis of organosolv poplar lignin by using high-performance liquid chromatography/high-resolution multi-stage tandem mass spectrometry. *Green Chemistry* **2021**, *23* (2), 983-1000.

(90) Sheng, H.; Tang, W.; Gao, J.; Riedeman, J.; Hurt, M.; Yang, L.; Kenttämaa, H. I. Characterization of ionized lignin model compounds with α-O-4 linkages by positive-and negative-ion mode electrospray ionization tandem mass spectrometry based on collision-activated dissociation. *Rapid Communications in Mass Spectrometry* **2021**, *35* (8), e9057.

(91) Qi, Y.; Fu, P.; Li, S.; Ma, C.; Liu, C.; Volmer, D. A. Assessment of molecular diversity of lignin products by various ionization techniques and high-resolution mass spectrometry. *Science of The Total Environment* **2020**, *713*, 136573.

(92) Dorrani, M.; Lynn, B. C. Application of lithium cationization tandem mass spectrometry for structural analysis of lignin model oligomers with β-β′ and β-O-4′ linkages. *Analytical and Bioanalytical Chemistry* **2022**, *414* (19), 5755-5771.

(93) Dean, K. R.; Lynn, B. C. Monolignol lithium cation basicity estimates and lithium adduct ion optimized geometries. *International Journal of Mass Spectrometry* **2019**, *442*, 109-116.

(94) Lancefield, C. S.; Westwood, N. J. The synthesis and analysis of advanced lignin model polymers. *Green Chemistry* **2015**, *17* (11), 4980-4990.

(95) NMR Database of Lignin and Cell Wall Model Compounds. https://www.glbrc.org/databases_and_software/nmrdatabase/NMR_DataBase_2009_Intro _and_Structure-Index.pdf (accessed 03/12/2023).

(96) Pavia, D. L.; Lampman, G. M.; Kriz, G. S.; Vyvyan, J. A. *Introduction to spectroscopy*; Cengage learning, 2014.

(97) Liu, W.; Liu, J.; Zhu, M.; Wang, W.; Wang, L.; Xie, S.; Wang, L.; Yang, X.; He, X.; Sun, Y. Recycling of lignin and Si waste for advanced Si/C battery anodes. *ACS Applied Materials & Interfaces* **2020**, *12* (51), 57055-57063.

(98) Falah, F.; Lubis, M. A. R.; Triastuti, T.; Fatriasari, W.; Sari, F. P. Utilization of lignin from the waste of bioethanol production as a mortar additive. *Jurnal Sylva Lestari* **2020**, *8* (3), 326-339.

(99) Liu, W.; Zhou, R.; Goh, H. L. S.; Huang, S.; Lu, X. From waste to functional additive: toughening epoxy resin with lignin. *ACS applied materials & interfaces* **2014**, *6* (8), 5810-5817.

(100) Sun, R. C. Lignin source and structural characterization. *ChemSusChem* **2020**, *13* (17), 4385-4393.

(101) Yang, J.; Ching, Y. C.; Chuah, C. H. Applications of lignocellulosic fibers and lignin in bioplastics: A review. *Polymers* **2019**, *11* (5), 751.

(102) Zhang, Y.; Liao, J.; Fang, X.; Bai, F.; Qiao, K.; Wang, L. Renewable high-performance polyurethane bioplastics derived from lignin–poly (ε-caprolactone). *ACS Sustainable Chemistry & Engineering* **2017**, *5* (5), 4276-4284.

(103) Witzler, M.; Alzagameem, A.; Bergs, M.; El Khaldi-Hansen, B.; Klein, S. E.; Hielscher, D.; Kamm, B.; Kreyenschmidt, J.; Tobiasch, E.; Schulze, M. Lignin-derived biomaterials for drug release and tissue engineering. *Molecules* **2018**, *23* (8), 1885.

(104) Dunky, M. Wood Adhesives Based on Natural Resources: A Critical Review: Part III. Tannin-and Lignin-Based Adhesives. *Progress in Adhesion and Adhesives* **2021**, *6*, 383-529.

(105) Lubis, M. A. R.; Labib, A.; Akbar, F.; Nuryawan, A.; Antov, P.; Kristak, L.; Papadopoulos, A. N.; Pizzi, A. Influence of lignin content and pressing time on plywood properties bonded with cold-setting adhesive based on poly (vinyl alcohol), lignin, and hexamine. *Polymers* **2022**, *14* (10), 2111.

(106) Tu, W.-C.; Hallett, J. P. Recent advances in the pretreatment of lignocellulosic biomass. *Current opinion in green and sustainable chemistry* **2019**, *20*, 11-17.

(107) Yu, J.; Zhang, J.; He, J.; Liu, Z.; Yu, Z. Combinations of mild physical or chemical pretreatment with biological pretreatment for enzymatic hydrolysis of rice hull. *Bioresource technology* **2009**, *100* (2), 903-908.

(108) Saritha, M.; Arora, A. Biological pretreatment of lignocellulosic substrates for enhanced delignification and enzymatic digestibility. *Indian journal of microbiology* **2012**, *52*, 122-130.

(109) Islam, M. K.; Rehman, S.; Guan, J.; Lau, C.-Y.; Tse, H.-Y.; Yeung, C. S.; Leu, S.-Y. Biphasic pretreatment for energy and carbon efficient conversion of lignocellulose into bioenergy and reactive lignin. *Applied Energy* **2021**, *303*, 117653.

(110) Marques, F. P.; Colares, A. S.; Cavalcante, M. N.; Almeida, J. S.; Lomonaco, D.; Silva, L. M.; de Freitas Rosa, M.; Leitão, R. C. Optimization by Response Surface

Methodology of Ethanosolv Lignin Recovery from Coconut Fiber, Oil Palm Mesocarp Fiber, and Sugarcane Bagasse. *Industrial & Engineering Chemistry Research* **2022**, *61* (11), 4058-4067.

(111) Hernández, Y. R.; García Serrano, L. A.; Maruri, D. T.; Jiménez Aparicio, A. R.; Camacho Díaz, B. H.; Arenas Ocampo, M. L. a. Optimization of the microwave-assisted ethanosolv extraction of lignocellulosic compounds from the bagasse of Agave angustifolia Haw using the response methodology. *Journal of agricultural and food chemistry* **2018**, *66* (13), 3533-3540.

(112) Kim, Y.; Yu, A.; Han, M.; Choi, G. W.; Chung, B. Ethanosolv pretreatment of barley straw with iron (III) chloride for enzymatic saccharification. *Journal of Chemical Technology & Biotechnology* **2010**, *85* (11), 1494-1498.

(113) Matsakas, L.; Raghavendran, V.; Yakimenko, O.; Persson, G.; Olsson, E.; Rova, U.; Olsson, L.; Christakopoulos, P. Lignin-first biomass fractionation using a hybrid organosolv–Steam explosion pretreatment technology improves the saccharification and fermentability of spruce biomass. *Bioresource technology* **2019**, *273*, 521-528.

(114) Wang, H.; Tang, X.; Arvanitis, M. A.; Yang, V.; Stark, N.; Liu, C.; Considine, J. M.; Zhu, J. Colloidal lignin nanoparticles from acid hydrotropic fractionation for producing tough, biodegradable, and UV blocking PVA nanocomposite. *Industrial Crops and Products* **2021**, *168*, 113584.

(115) Lu, F.; Ralph, J. DFRC method for lignin analysis. 1. New method for β-aryl ether cleavage: lignin model studies. *Journal of Agricultural and Food Chemistry* **1997**, *45* (12), 4655-4660.

(116) Lu, F.; Ralph, J. The DFRC (Derivatization Followed by Reductive Cleavage) method and its applications for lignin characterization. *Lignin: structural analysis, applications in biomaterials, and ecological significance* **2014**, 27-65.

(117) Ji, H.; Wang, L.; Tao, F.; Yao, Z.; Li, X.; Dong, C.; Pang, Z. A hydrotrope pretreatment for stabilized lignin extraction and high titer ethanol production. *Bioresources and Bioprocessing* **2022**, *9* (1), 1-11.

(118) Meng, X.; Bhagia, S.; Wang, Y.; Zhou, Y.; Pu, Y.; Dunlap, J. R.; Shuai, L.; Ragauskas, A. J.; Yoo, C. G. Effects of the advanced organosolv pretreatment strategies on structural properties of woody biomass. *Industrial Crops and Products* **2020**, *146*, 112144.

(119) Patri, A. S.; Mohan, R.; Pu, Y.; Yoo, C. G.; Ragauskas, A. J.; Kumar, R.; Kisailus, D.; Cai, C. M.; Wyman, C. E. THF co-solvent pretreatment prevents lignin redeposition from interfering with enzymes yielding prolonged cellulase activity. *Biotechnology for Biofuels* **2021**, *14*, 1-13.

(120) Zhao, Z.-M.; Meng, X.; Scheidemantle, B.; Pu, Y.; Liu, Z.-H.; Li, B.-Z.; Wyman, C. E.; Cai, C. M.; Ragauskas, A. J. Cosolvent enhanced lignocellulosic fractionation tailoring lignin chemistry and enhancing lignin bioconversion. *Bioresource Technology* **2022**, *347*, 126367.

(121) Ma, Q.; Hirth, K.; Zhai, H.; Zhu, J. Highly bleachable wood fibers containing less condensed lignin from acid hydrotropic fractionation (AHF). *ACS Sustainable Chemistry & Engineering* **2020**, *8* (24), 9046-9057.

(122) Schäfer, J.; Urbat, F.; Rund, K.; Bunzel, M. A stable-isotope dilution GC-MS approach for the analysis of DFRC (derivatization followed by reductive cleavage)

monomers from low-lignin plant materials. *Journal of agricultural and food chemistry* **2015**, *63* (10), 2668-2673.

(123) Zhang, L.; Larsson, A.; Moldin, A.; Edlund, U. Comparison of lignin distribution, structure, and morphology in wheat straw and wood. *Industrial Crops and Products* **2022**, *187*, 115432.

Poorya Kamali was born in Tehran, Iran. He obtained his B.Sc in Applied Chemistry from University of Mohaghegh Ardabili, Ardabil, IR. He joined University of Kentucky to persue his graduate studies under Dr. Bert C. Lynn's supervision and earned his Ph.D. degree in Analytical Chemistry in 2023.

Award/Honors:

- Best presenter in National Chemistry and Nanotechnology Conference, 2014
- C.H.H. Griffith Outstanding General Chemistry Teaching Assistant Award, 2023

Publications:

- Asare, S. O., **Kamali, P**., Huang, F., & Lynn, B. C. (2018). Application of chloride adduct ionization tandem mass spectrometry for characterizing and sequencing synthetic lignin model compounds. *Energy & fuels*, 32(5), 5990-5998.

- Tong, X., Moradipour, M., Novak, B., **Kamali, P**., Asare, S. O., Knutson, B. L., Rankin, S. E., Lynn, B. C. & Moldovan, D. (2019). Experimental and molecular dynamics simulation study of the effects of lignin dimers on the gel-to-fluid phase transition in DPPC bilayers. *The Journal of Physical Chemistry B*, 123(39), 8247-8260.

- Moradipour, M., Tong, X., Novak, B., **Kamali, P**., Asare, S. O., Lynn, B. C., Moldovan, D., Rankin, S. E., & Knutson, B. L. (2021). Interaction of lignin dimers with model cell membranes: A quartz crystal microbalance and molecular dynamics simulation study. *Biointerphases*, 16(4), 041003.

- Dorrani, M., **Kamali, P**., Sunder, P. R., Dean, K. R., Sadik, A., Lynn, B. C. A universal GC-MS method for the analysis of lignin-derived products. *In-progress*

Selected Presentations:

- **Kamali P**., Lynn B.C., "Application of gas chromatography -mass spectrometry for the analysis of structural isomers of lignin dimers", *American Society of Mass Spectrometry*, Atlanta, USA, 2019.

- **Kamali, P**., Asare, S., Lynn, B. "Synthesis and characterization of β-O-4 trimer precursor and putative synthesis of a novel β-O-4/β-5 lignin trimer model compound" *2018 Annual Louisiana and Kentucky Interjurisdictional Conference*, Baton Rouge, LA. USA.

- **Kamali, P**., Asare, S., Lynn, B. "Identification of an unusual by-product observed in lignin precursor synthesis" *2017 Annual Louisiana and Kentucky Interjurisdictional Conference*, Lexington, KY. USA