APPLICATION OF HIGH-RESOLUTION ACCURATE MASS (HRAM) MASS SPECTROMETRY FOR ANALYSIS OF LIGNIN MODEL COMPOUNDS AND THE POST-PRETREATMENT PRODUCTS

Fan Huang

University of Kentucky, fhuang529@gmail.com
Digital Object Identifier: https://doi.org/10.13023/ETD.2017.149

Click here to let us know how access to this document benefits you.

Recommended Citation
https://uknowledge.uky.edu/chemistry_etds/74

This Doctoral Dissertation is brought to you for free and open access by the Chemistry at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Chemistry by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
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.

Fan Huang, Student

Dr. Bert C. Lynn, Major Professor

Dr. Mark A. Lovell, Director of Graduate Studies
APPLICATION OF HIGH-RESOLUTION ACCURATE MASS (HRAM) MASS SPECTROMETRY FOR ANALYSIS OF LIGNIN MODEL COMPOUNDS AND THE POST-PRETREATMENT PRODUCTS

DISSERTATION

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

By

Fan Huang

Lexington, Kentucky

Director: Dr. Bert C Lynn, Professor of Chemistry

Lexington, Kentucky

2017

Copyright © Fan Huang 2017
ABSTRACT OF DISSERTATION

APPLICATION OF HIGH-RESOLUTION ACCURATE MASS (HRAM) MASS SPECTROMETRY FOR ANALYSIS OF LIGNIN MODEL COMPOUNDS AND THE POST-PRETREATMENT PRODUCTS

Lignin, one of the main components in the woody cell walls, is a complex heterogeneous biopolymer, which provides structural support and transportation of water in plants. It is highly recalcitrant to degradation (both chemically and environmentally) and protects cellulose from being degraded/hydrolyzed. Due to the structural complexity of native lignin, complete characterization and elucidation of lignin’s structure remains very challenging. The overarching goal of this work is to develop mass spectrometry-based analytical methods to contribute to a better understanding of lignin structures.

This dissertation will focus on the development and application of High-Resolution Accurate-Mass (HRAM) Mass Spectrometry (MS) as the main analytical technique for studying lignin model compounds, including understanding the ionization behavior, studying corresponding fragmentation patterns and extracting structural information for structural elucidation eventually. Analytical methods were also developed to study the post-pretreatment products of the synthetic trimeric model compound using High-Performance Liquid Chromatography (HPLC) coupled with High-Resolution Accurate Mass (HRAM) Mass Spectrometry (MS).

The first project of this dissertation focuses on mass spectral characterization of lignin models from the in-vitro oxidative coupling reactions. Three specific trimeric compounds were isolated and their ionization behaviors were investigated using HRAM-MS via electrospray ionization (ESI). The reaction parameters of the in vitro oxidative coupling reaction were critical in selecting the linkage profiles of resulting dehydrogenation polymers (DHPs). Reaction parameters were tuned to obtain desired DHP linkages profile. Upon the isolation of three different trimeric compounds, a systematic comparison of ionization efficiency of three trimeric compounds was carried out using ESI-HRAM-MS under different ionization conditions.

The second project was aimed to design a synthetic route for a lignin model compound that will be a good representation for native lignin during the pretreatment process. The model compound of interest has not been obtained previously through chemical synthesis. Due to the reactivity of cinnamyl alcohol, which contains the unsaturated side chain, this new synthesis strategy was developed based on the known
aldol-type reaction route. A versatile synthesis procedure for preparation of β-O-4 oligomeric compounds was designed and implemented to include the most important functional groups (phenolic alcohol, aryl glycerol β-aryl ether bond and unsaturated side chain) in the resulting model compound. This new synthesis route also allowed incorporation of different monolignols.

In the third project, Fenton chemistry was applied to a synthetic lignin model compound. Due to the non-specificity in the post pretreatment product profile, non-targeted analytical strategy was developed and applied to study the post-pretreatment products of the model compound using HPLC-HRMS.

The results from this dissertation showed a significant difference in ionization behavior between three structurally different model compounds and indicated that primary structures of lignin compounds can largely affect corresponding electrospray ionization properties as well as fragmentation pattern. The work in this dissertation provides analytical techniques for non-targeted analysis of complex lignin samples and an insightful understanding of Fenton’s reaction pretreatment upon lignin model compound.

KEYWORDS: High-Resolution Mass Spectrometry, Lignin, Biofuel, Liquid Chromatography, Electrospray Ionization

________________________________________
Fan Huang

4/25/2017
APPLICATION OF HIGH-RESOLUTION ACCURATE MASS (HRAM) MASS SPECTROMETRY FOR ANALYSIS OF LIGNIN MODEL COMPOUNDS AND THE POST-PRETREATMENT PRODUCTS

By

Fan Huang

Bert C. Lynn

(Director of Dissertation)

Mark A. Lovell

(Director of Graduate Studies)

4/25/2017
Dedicated to my friends and family who supported me through this journey
ACKNOWLEDGEMENTS

I would like to thank my parents for their unconditional love and trust for me every moment of my life. Their love and support have carried me throughout my whole graduate life. It is such a blessing for me to grow up in such a supportive and loving family, which has grown me into who I am today.

All these would not happen without my advisor, Dr. Bert C. Lynn, who deserves my sincerest gratitude. He has been giving me invaluable support and mentoring during this journey in seeking my PhD degree. I am always amazed by how much and how diverse his knowledge is. It was definitely a wonderful and incredible experience to learn from him. He has taught me to take initiatives, to think independent and to apply critical-thinking in becoming a scientist. I am so grateful for him being such an open-minded and patient advisor whenever I overwhelm him with hundreds of “why’s”. Not only has he given me professional guidance throughout the years in his group, but he has also supported me and given me invaluable advice regarding my future career decisions. Dr. Lynn is the best advisor and mentor I could ever have. I would also like to sincerely thank all my committee members, Dr. Yinan Wei, Dr. Jason DeRouchey and Dr. Robert Lodder for all their supports and insightful guidance in helping me throughout my graduate years. I would like to thank Dr. Michael Flythe for being allowing me to work in his lab and teaching me skills to work with bacteria in the beginning of this journey, it has been an enjoyable experience to work with him. I would like to thank Dr. Sue Nokes for serving as my outside examiner; it has also been a pleasant experience to collaborate with her. She has also been very supportive. I would like to express my gratitude to John Layton and Dr. Anne-Frances Miller for providing me incredible guidance in running
NMR experiments. I would also like to thank Dr. Sean Parkin for helping me obtaining X-ray crystal structure data whenever I ask for.

I want to thank all the past and present group members in Dr. Lynn’s group. I would like to specifically thank Jennifer Ashley, Dr. Dawn Kato, Dr. Brent Casper and Dr. Sanja Trajkovic for their supporting from the beginning when I joined Dr. Lynn’s group. I have learned many diverse analytical skills from their experience which allowed me to quickly adapt into the life of graduate school. I am grateful to have them around me during the highs and lows throughout the process, since they understand this journey the best. I also want to express my gratitude to all my organic chemistry friends in the Department of Chemistry, who have been my resources whenever I have organic questions.

Last but certainly not the least, I would also like to thank all my friends, both in China and here in the States. They have been encouraging and supporting me through this journey. I am so grateful to know all these wonderful people during my time at University of Kentucky. Their love and support have made me feel as if I am home even though I am on the other side of the earth.
# TABLE OF CONTENTS

ACKNOWLEDGEMENT ........................................................................................................ iii

LIST OF TABLES .................................................................................................................. vii

LIST OF FIGURES ............................................................................................................... viii

LIST OF ABBREVIATIONS AND SYMBOLS ........................................................................ xi

Chapter 1: Introduction ....................................................................................................... 1

1.1 Background .................................................................................................................. 1

1.2 Lignocellulosic Biomass ............................................................................................... 6

1.3 Conversion of Lignocellulosic Biomass to Biofuels ...................................................... 10

1.4 High-Performance Liquid Chromatography ................................................................. 13

1.5 Electrospray Ionization-Orbitrap Mass Analyzer ......................................................... 14

Chapter 2: Ionization Behavior of Lignin Model Compounds ........................................... 24

2.1 Introduction ................................................................................................................ 24

2.1.1 Formation and Structure of Lignin ........................................................................ 24

2.1.2 Dehydrogenation Polymers (DHPs) .................................................................... 25

2.1.3 Methods for Characterization of Lignin ................................................................. 31

2.2 Materials and Methods ............................................................................................... 39

2.3 Results and Discussion ............................................................................................... 46

2.3.1 Characterization of Two Different DHPs ................................................................. 46

2.3.2 Negative Ionization Behavior of Trilignols ............................................................. 53

2.3.3 Positive Ionization Behavior of Trilignols ............................................................... 60

2.3.4 HCD/MS/MS Experiment of Trilignols ................................................................. 67

2.4 Conclusion .................................................................................................................. 77

Chapter 3: Synthesis of Trimeric Lignin Model Compounds ............................................. 78

3.1 Introduction ................................................................................................................ 78

3.1.1 Lignin Model Compounds ..................................................................................... 78

3.1.2 Condensation Reaction of Aldehydes and Ketones .................................................. 82

3.2 Materials and Methods ............................................................................................. 84

3.3 Results and Discussion ............................................................................................. 91
LIST OF TABLES

Table 2.1  A brief review of different chemical degradation methods for structural characterization of native lignin…………………………………………………32

Table 2.2  NMR chemical shift assignments of three trimers…………………………………….45

Table 2.3  List of observed fragment ions for G(β-O-4)G(β-5)G and their chemical formula based on high resolution MS result of [M+Li]^+ and [M+Na]^+…73

Table 2.4  List of observed fragment ions for G(β-O-4)G(β-β) and their chemical formula based on high resolution MS result of [M+Li]^+ and [M+Na]^+....76

Table 4.1  Summary of various pretreatment strategies………………………………………..103

Table 4.2  LC method used in the pretreatment study………………………………………..111

Table 4.3  List of 20 features with their corresponding elemental composition and retention time based on the result from MZmine 2…………………..116
LIST OF FIGURES

Figure 1.1 Total Fossil Fuels Productions and primary energy production by source (Data collected by EIA, Annual Energy Outlook, 2015) ....................................3
Figure 1.2 Total Renewable Energy Productions and primary energy production by sources (Data collected by EIA, Annual Energy Outlook, 2015) ..........4
Figure 1.3 Primary Energy Consumption by source in 2015 (Data collected by EIA, Annual Energy Outlook, 2015) ..........................................................5
Figure 1.4 Simplified schematics of lignocellulose structure ...........................................8
Figure 1.5 Common linkage types presented in lignin .........................................................9
Figure 1.6 Typical bioconversion process of biomass-to-biofuel ........................................12
Figure 1.7 Schematic depiction of an ESI source operated in the positive ion mode ...........................................................................................................17
Figure 1.8 Simplified Schematics of an Orbitrap .................................................................20
Figure 1.9 Schematic of Q-Exactive mass spectrometer (Thermo Scientific) (source: http://planetorbitrap.com/q-exactive#tab:schematic) .............................................23
Figure 2.1 Structures of three monolignols: a) coniferyl alcohol(G); b) p-coumaryl alcohol(H); c) sinapyl alcohol(S) .................................................................26
Figure 2.2 Resonance structure of phenoxy radical of lignin monomer ................................27
Figure 2.3 Three Examples of common bonding types presented in lignin; a) β-aryl ether (β-O-4); b) resinol (β-β); c) phenylcoumaran (β-5) ............................28
Figure 2.4 Example of β-O-4 linkage formation via radical coupling .................................29
Figure 2.5 Structures of three trimers: a) Guaiacylglycerol-β, γ-bis-coniferyl ether [G(4-O-α)G(β-O-4)G]; b) Guaiacylglycerol-β-dehydrodiconiferyl ether [G(β-O-4)G(β-5)G]; c) Guaiacylglycerol-pinoresinol ether [G(β-O-4)G(β-β)G] .........................................................44
Figure 2.6 Full scan mass spectrum of DHP in Positive ion mode ESI, [M+NH₄]⁺; a) DHP-CT; b) DHP-ZT .........................................................................................47
Figure 2.7 Full scan mass spectrum of DHP in Negative ion mode ESI, [M+NH₄]⁺; a) DHP-CT; b) DHP-ZT .............48
Figure 2.8 GPC chromatogram of isolated DHPs; from top to bottom were coniferyl alcohol, dimers, trimers, tetramers and pentamers ..........................51
Figure 2.9 GPC chromatogram of DHPs from two different reaction conditions: DHP-CT products (top) and DHP-ZT products (bottom) ................ 52
Figure 2.10 a) Deprotonated ion intensity comparison of three trimers; b) full scan mass spectrum of G(4-O-α)G(β-O-4)G, Error bars indicate ± one standard deviation .........................................................55
Figure 2.11 Liquid chromatograms of G(4-O-α)G(β-O-4)G, a) BPC of trimer 537.213; b) EIC of 537.213; c) EIC of dimer, m/z 357.134; d) EIC of monomer, m/z 179.071; e) EIC of coniferyl alcohol, 179.071.

Figure 2.12 Proposed fragmentation mechanism of deprotonated G(4-O-α)G(β-O-4)G ion.

Figure 2.13 Ion intensity comparison of three trimers with chlorine ion adduct, [M+Cl]− in negative ion mode ESI, error bars indicate ± one standard deviation.

Figure 2.14 a) Relative ionization response of G(4-O-α)G(β-O-4)G in different salt solution. b) Cation selectivity of G(4-O-α)G(β-O-4)G in solution of equal molar of NH4Cl/LiCl/NaCl/KCl/RbCl, error bars indicate ± one standard deviation.

Figure 2.15 a) Relative ionization response of G(β-O-4)G(β-5)G in different salt solution. b) Cation selectivity of G(β-O-4)G(β-5)G in solution of equal molar of NH4Cl/LiCl/NaCl/KCl/RbCl, error bars indicate ± one standard deviation.

Figure 2.16 a) Relative ionization response of G(β-O-4)G(β-β)G in different salt solution. b) Cation selectivity of G(β-O-4)G(β-β)G in solution of equal molar of NH4Cl/LiCl/NaCl/KCl/RbCl, error bars indicate ± one standard deviation.

Figure 2.17 Positive ion mode ESI response comparison of three trilignols with the presence of different cations, error bars indicate ± one standard deviation.

Figure 2.18 HCD/MS/MS spectra of three trimers with ammonium adduct, [M+NH4]+, a) G(4-O-α)G(β-O-4)G; b) G(β-O-4)G(β-5)G; c) G(β-O-4)G(β-β)G.

Figure 2.19 HCD/MS/MS spectra of G(4-O-α)G(β-O-4)G: a) [M-H]; b) [M+Li]+; c) [M+Na]+.

Figure 2.20 HCD/MS/MS spectra of G(β-O-4)G(β-5)G: a) [M-H]; b) [M+Li]+; c) [M+Na]+.

Figure 2.21 HCD/MS/MS spectra of G(β-O-4)G(β-β)G: a) [M-H]; b) [M+Li]+; c) [M+Na]+.

Figure 3.1 Synthetic route of target trimer G(β-O-4)G(β-O-4)G.

Figure 3.2 General mechanism of a typical aldol reaction.

Figure 3.3 The synthesis of the three building blocks for the synthesis of target trimer.

Figure 3.4 X-ray crystal structure of compound 9.

Figure 3.5 a) Reaction pathway of nucleophilic addition of di-isopropanol amine to compound 6 (purple color); nucleophilic addition within compound 6.
Figure 3.6  
(a) ESI-MS of targeted trimer in negative ion mode, [M-H];  
(b) HCD/MS/MS spectrum of the trimer, [M-H]  

Figure 4.1  
Schematic of pretreatment step of biomass 

Figure 4.2  
Liquid chromatogram of mixture of standards;  
(a) Base peak chromatogram (BPC);  
(b) Extracted Ion Chromatogram (EIC) of 3-hydroxy-4-methoxybenzoic acid;  
(c) EIC of ferulic acid;  
(d) EIC of coniferyl alcohol;  
(e) EIC of vanillin 

Figure 4.3  
EIC of starting trilignol GGG 

Figure 4.4  
Time course measurement of starting material GGG, error bars indicate ± one standard deviation 

Figure 4.5  
Degradation of trimeric unknown feature, [M-H]  
551.1909 

Figure 4.6  
(a) HCD/MS/MS spectrum of m/z 551.1909;  
(b) putative structures of m/z 551.1909 

Figure 4.7  
possible fragmentation structures for corresponding trimeric unknown  
m/z 551.1909 

Figure 4.8  
Time course plots of fifteen unknown features, error bars indicate ± one standard deviation 

Figure 4.9  
Extract ion chromatogram (EIC) and HCD/MS/MS spectrum of vanillin (upper panel), EIC and HCD/MS/MS spectrum of unknown m/z 151.0397 (lower panel) 

Figure 4.10  
EIC and HCD/MS/MS spectrum of vanillic acid (upper panel), EIC and HCD/MS/MS spectrum of m/z 167.0346 (lower panel) 

Figure 4.11  
EIC and HCD/MS/MS spectrum of synthetic G(β-O-4)G(β-O-4)G (upper panel), EIC and HCD/MS/MS spectrum of m/z 375.1439 (lower panel) 

Figure 4.12  
HCD/MS/MS spectra of three unknown features with putative structures
# LIST OF ABBREVIATION AND SYMBOLS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFEX</td>
<td>Ammonium fiber expansion</td>
</tr>
<tr>
<td>APCI</td>
<td>Atmospheric Pressure Chemical Ionization</td>
</tr>
<tr>
<td>AU</td>
<td>Arbitrary Units</td>
</tr>
<tr>
<td>BPC</td>
<td>Base Peak Chromatogram</td>
</tr>
<tr>
<td>Btu</td>
<td>British thermal unit</td>
</tr>
<tr>
<td>°C</td>
<td>Degree Celsius</td>
</tr>
<tr>
<td>C-trap</td>
<td>Curved quadrupole ion trap</td>
</tr>
<tr>
<td>C18</td>
<td>octadecyl carbon chain</td>
</tr>
<tr>
<td>CAD</td>
<td>Collision Activated Dissociation</td>
</tr>
<tr>
<td>CID</td>
<td>Collision induce dissociation</td>
</tr>
<tr>
<td>COSY</td>
<td>Homonuclear correlation spectroscopy</td>
</tr>
<tr>
<td>(CTA)$_2$SO$_4$</td>
<td>Cetyltrimethylammonium sulfate</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DFRC</td>
<td>Derivatization Followed by Reductive Cleavage</td>
</tr>
<tr>
<td>DHP-CT</td>
<td>Dehydrogenation Polymer- Cetyltrimethylammonium sulfate</td>
</tr>
<tr>
<td>DHP-ZT</td>
<td>Dehydrogenation Polymer- Zutropf</td>
</tr>
<tr>
<td>DHPs</td>
<td>Dehydrogenation Polymers</td>
</tr>
<tr>
<td>DIBAL-H</td>
<td>Diisobutylaluminum Hydride</td>
</tr>
<tr>
<td>Dpmo</td>
<td>Number Degree of Polymerization</td>
</tr>
<tr>
<td>EI</td>
<td>Electron Ionization</td>
</tr>
<tr>
<td>EIA</td>
<td>Energy Information Administration</td>
</tr>
<tr>
<td>EIC</td>
<td>Extracted Ion Chromatogram</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray Ionization</td>
</tr>
<tr>
<td>ESI-MS</td>
<td>Electrospray Ionization-Mass Spectrometry</td>
</tr>
<tr>
<td>FT-ICR</td>
<td>Fourier Transform Ion Cyclotron Resonance</td>
</tr>
<tr>
<td>G</td>
<td>Guaiacyl</td>
</tr>
<tr>
<td>GC</td>
<td>Gas Chromatography</td>
</tr>
<tr>
<td>GC/MS</td>
<td>Gas Chromatography/Mass Spectrometry</td>
</tr>
<tr>
<td>GGG</td>
<td>G(β-O-4)G(β-O-4)G</td>
</tr>
<tr>
<td>GOH</td>
<td>Coniferyl Alcohol</td>
</tr>
<tr>
<td>GPC</td>
<td>Gel Permeation Chromatography</td>
</tr>
<tr>
<td>H</td>
<td>Hydroxyphenyl</td>
</tr>
<tr>
<td>HCD</td>
<td>High Energy Collision Cell</td>
</tr>
<tr>
<td>HCD/MS/MS</td>
<td>High-Energy Collision Dissociation/Tandem Mass Spectrometry</td>
</tr>
<tr>
<td>HPLC</td>
<td>High-Performance Liquid Chromatography</td>
</tr>
<tr>
<td>HRAM</td>
<td>High-Resolution Accurate Mass</td>
</tr>
<tr>
<td>HRMS</td>
<td>High-Resolution Mass Spectrometry</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear Single Quantum Correlation</td>
</tr>
<tr>
<td>IR</td>
<td>Infrared</td>
</tr>
<tr>
<td>kV</td>
<td>Kilovolt</td>
</tr>
<tr>
<td>LC</td>
<td>Liquid Chromatography</td>
</tr>
<tr>
<td>LC-MS</td>
<td>Liquid Chromatography-Mass Spectrometry</td>
</tr>
<tr>
<td>LDA</td>
<td>Lithium Diisopropylamide</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass-to-charge ratio</td>
</tr>
<tr>
<td>MALDI</td>
<td>Matrix Laser Desorption Assisted Ionization</td>
</tr>
<tr>
<td>MALDI-MS</td>
<td>Matrix Assisted Laser Desorption Ionization-Mass Spectrometry</td>
</tr>
<tr>
<td>MS</td>
<td>Mass Spectrometry</td>
</tr>
<tr>
<td>MS/MS</td>
<td>Tandem Mass Spectrometry</td>
</tr>
<tr>
<td>MSW</td>
<td>Municipal Solid Waste</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular Weight</td>
</tr>
<tr>
<td>MWL</td>
<td>Mill Wood Lignin</td>
</tr>
<tr>
<td>NaHMDS</td>
<td>Sodium Bis(trimethylsilyl)amide</td>
</tr>
<tr>
<td>NIR</td>
<td>Near Infrared</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>ppm</td>
<td>Parts Per Million</td>
</tr>
<tr>
<td>PPP</td>
<td>Para-phenylphenol</td>
</tr>
<tr>
<td>RP</td>
<td>Reversed-phase</td>
</tr>
<tr>
<td>rt</td>
<td>Retention Time</td>
</tr>
<tr>
<td>S</td>
<td>Syringyl</td>
</tr>
<tr>
<td>SEC</td>
<td>Size Exclusion Chromatography</td>
</tr>
<tr>
<td>STDEV</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>TBDMS</td>
<td>t-butyldimethylsilyl</td>
</tr>
<tr>
<td>TIPS</td>
<td>Triisopropylsilyl</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin Layer Chromatography</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>ZL</td>
<td>Zulauf</td>
</tr>
<tr>
<td>ZT</td>
<td>Zutropf</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Background

According to U.S. Energy Information Administration (EIA), combined energy consumption rose from a low of 32 quadrillion Btu to a high of 97 quadrillion Btu since 1949 to 2015\(^1-2\). This is an approximate 200% increase in energy consumption since 1949. There has been a downward trend in the production and consumption of coal and petroleum since 2006. The global energy production highly depends on the production of fossil fuels, which are the major energy sources. As global energy production and consumption have been growing increasingly, it may eventually lead to the depletion of fossil fuels. Alternative energy resources that meet the requirements of being renewable, sustainable and ecofriendly are the fundamental solution to the current concerns regarding both energy sustainability as well as global environmental problems created from conventional energy consumption. Solar energy is the most sustainable and abundant form of renewable energy on earth. Currently, solar is being utilizing in four ways: 1) direct photovoltaic conversion of sunlight to electricity; 2) via wind turbines to produce electricity; 3) by using mirrors to concentrate solar to power stirling engine which produces electricity; 4) by harvesting plants that can be burned to generate energy\(^3\). Among these different methods, plants are the most cost-effective one in converting photons into chemically stored energy\(^4\). Although the production of biomass, wind, solar and other renewable energies are trending upward, the production of natural gas has grown the most in the past ten years. As shown in Figure 1.1-1.3, according to the data collected by EIA, there is an upward trend of total energy production from 1949 to
Natural gas and crude oil production are accounted for the increasing trend of total fossil fuels production. Nuclear electric power and biomass-based energy are two major sources accounted for total renewable energy production.

The second-largest biomass uses are wood and wood waste for production of electric and industrial power. For liquid transportation fuels production, current primary industrial biomass resources are predominantly corn grain for bioethanol and soybean oil for biodiesel. In general, technologies that convert corn grain and soybean oil to liquid fuel are referred as “first generation”\(^2\). However, the first-generation biomass, such as corn and soybean, are also important food resources. A concern has been that the food based-biofuel production contributes significantly increased in food prices and in the agricultural cultivation cost and land concentration for food based biomass, which results in the increase of total cost for food-based biofuel production. There is also an ethical concern that conventional biofuel production might reduce access to food and increase the food price, ultimately increasing global hunger.

Advanced technologies now enable the production of biofuel from cellulosic biomass, which is referred as second-generation, non-grain, and non-food based feedstock. There are several types of cellulosic biomass: 1) agricultural residues, which are non-food based by-products (e.g. corn stover, cotton gin trash); 2) energy crops, including woody energy crops (e.g., hybrid poplars, shrub willows) and herbaceous energy crops (e.g., switchgrass, miscanthus, sorghum, energycane); 3) forest resources: wood waste and woody biomass, including existing and re-purposed pulp and paper products, logging residues and forest thinning; 4) industrial and other waste, such as municipal solid waste (MSW) and urban renewal wood;
Figure 1.1. Total Fossil Fuels Productions and Primary Energy Production by source

(Data collected by EIA, Annual Energy Outlook, 2015)
Figure 1.2. Total Renewable Energy Productions and Primary Energy Production by sources (Data collected by EIA, Annual Energy Outlook, 2015).
Figure 1.3. Primary Energy Consumption by source in 2015 (Data collected by EIA, Annual Energy Outlook, 2015).
5) algae, which is a diverse group of primarily aquatic photosynthetic algae and cyanobacteria.

1.2 Lignocellulosic Biomass

Bioenergy plants generally fall into two categories: 1) gymnosperms, such as soft woods (e.g., pine, spruce and cedar); 2) angiosperms including perennial grasses (e.g., switchgrass, miscanthus and sorghum), herbaceous species (e.g., corn, wheat), flowering plants (e.g., alfalfa, soybean and tobacco) and hard woods (e.g., poplar, willow and black locust). The major component of plant biomass is lignocellulose, which forms the plant cell walls, supporting and providing plant a protective barrier from degradation by physiological environment. Lignocelluloses are mainly composed of cellulose, hemicellulose and lignin. These three components interact in very intricate manner, resulting in heavily recalcitrant cell wall that is difficult to degradation or decomposition (Figure 1.4).

Cellulose, the major component of lignocellulosic biomass (35-50%), is a linear polymer of several hundred to thousands of cellobiose units, which consists of two D-glucose molecules linked through β- (1,4)-glycosidic bonds. The cellulose strains are held together firmly through intra- and inter- molecular hydrogen bonding to form cellulose fibrils with high tensile strength. The cellulose fibers are further linked together via hydrogen bonds resulting in high rigidity and crystallinity of dominant cellulose. Therefore, cellulose is insoluble in water and most organic solvents.

Hemicelluloses (20-35%) are heterogeneous branched polymers that include several polysaccharides, such as arabinoxylan, xylan, glucuronoxylan, glucomannan and
xyloglucan. Hemicelluloses contain many different sugar monomers, for instance, xylose, mannose, galactose and arabinose. In contrast to crystallinity and rigidity of cellulose, hemicelluloses are easier to be hydrolyzed due to their amorphous, branched structure with short polysaccharide chain. However, hemicelluloses reduce the accessibility of cellulose protecting cellulose from hydrolysis.\(^6\)

Lignin (15-20%) is a class of complex aromatic biopolymers that plays an important role in the support of vascular plants. Lignin is a polymer with three basic phenylpropane units: \(p\)-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol. The content of lignin depends on the type of species. Lignin structures can also differ within the same plant, for instance, primary xylem, compression wood or early versus late wood cell. The monomeric composition in lignin also varies in different type of plants: \(p\)-coumaryl alcohol mainly exists in compression wood and grasses; coniferyl alcohol largest exists both in hardwoods and softwoods; sinapyl alcohol mostly presents in hardwoods. Biosynthesis of lignin is through radical oxidative coupling of three monolignols. It results in the extreme complex amorphous polymers that contain variety chemical linkages: phenylpropane-\(\beta\)-aryl ether (\(\beta\)-O-4), phenylcoumaran (\(\beta\)-5), resinol (\(\beta\)-\(\beta\)), biphenyl (5-5), dibenzodioxocin, 1,3-diarylpropane (\(\beta\)-1), \(\alpha\)-aryl ether (\(\alpha\)-O-4) and diaryl ether (4-O-5). The frequency of different linkage types differs between species (showed in Figure 1.5). \(\beta\)-aryl ether (\(\beta\)-O-4) is considered to be the most abundant linkages in lignin (~ 60% in hard wood, 50% in soft wood). Phenylcoumaran (\(\beta\)-5) is about 11% in soft wood and 6% in hard wood. \(\alpha\)-aryl ether (\(\alpha\)-O-4) is considered to be a relatively labile linkage in lignin compared with other linkages, it could present in dibenzodioxocin linkage or \(\alpha\), \(\beta\)-diaryl ether linkage.\(^7\)
Figure 1.4. Simplified schematics of lignocellulose structure.
β-aryl ether

Resinol

Phenylcoumaran

Dibenzodioxocin

α, β-diaryl ether

Biphenyl

Diaryl ether

Figure 1.5. Common linkage types presented in lignin.
1.3 Conversion of Lignocellulosic Biomass to Biofuels

The techniques for conversion of biomass to biofuels fall into two different methods: thermochemical and biochemical conversion.

Thermochemical conversion usually requires high-temperature deconstruction of biomass, such as pyrolysis, gasification and hydrothermal liquefaction. Pyrolysis decomposes biomass without the presence of oxygen to produce bio-oil intermediate at high temperature with or without the presence of catalysts. The resulting bio-oil contains mixtures of hydrocarbons with various lengths. The product requires further upgrading before it can be finished into a fuel or used in a refinery because it contains more oxygenated compounds than petroleum crude oils. Gasification typically deconstructs of biomass at more than 700 degree Celsius in the presence of sub-stoichiometric air generating clean synthesis gas or syngas, a mixture of CO₂, CO, CH₄, N₂, and H₂. The syngas can be further converted by catalyst to various other molecules such as methanol, ethanol and dimethyl ether. Hydrothermal liquefaction is a thermal deconstruction process of wet feedstock slurry under elevated temperature and pressure to generate hydrothermal liquefaction bio-oil. This process is mostly applicable to algal feedstock.³,⁸

Biochemical conversion techniques require much lower temperature compared with thermochemical conversion. In general, biochemical conversion includes pretreatment, enzymatic hydrolysis, fermentation and product recovery (Figure 1.6)⁸. Bioconversion technologies appear to be more attractive over thermochemical method for several reasons: 1) it is relatively scale-neutral and may have lower capital costs than thermal conversion; 2) it is possible that technical development can reduce the costs of
production of cellulosic fuels so that farmers can participate in the ownership of the facilities.

As described in previous section, due to the intricate biological structure of lignocellulosic feedstock, biomass remains naturally recalcitrant, protecting themselves from decomposition. Therefore, one of the key steps in bioconversion is pretreatment of biomass. The goal of pretreatment is to use biological or chemical methods to increase the porosity of biomass feedstock and degrade/remove/modify lignin to increase the accessibility of cellulose and other polysaccharides for enzymatic hydrolysis in the following step. During the enzymatic hydrolysis, the solubilized cellulose and other polysaccharides are subjected to further hydrolysis to simple sugars such as glucose. More efforts have been put on search for more active glycosyl hydrolases from incompletely explored sources (e.g. termites). The simple sugars are then fermented to liquid fuels such as ethanol, 1-butanol and acetone by the industrial biofuel-producing microorganism during the fermentation step. The generated liquid fuels are then recovered and separated from the fermentation step and ready for commercial purpose.

The performances of different pretreatment methods are typically evaluated by the level increased in the bioavailability of cellulose after the pretreatment and the level of total lignin content changes according to National Renewable Energy Laboratory (NREL) standard protocols. However, this method does not provide insights regarding the structural changes in lignin caused by the pretreatment process. Chemical degradation techniques coupled with gas chromatography (GC) have been applied in analysis of monomeric units of lignin. However, GC analysis has limitations on the size of the molecules.
Figure 1.6. Typical bioconversion process of biomass-to-biofuel
Typically, for analysis of lignin compounds, GC fails when the molecular weight higher than dimeric compounds.

High-Performance Liquid chromatography (HPLC) is commonly used to provide versatile separation abilities for complex sample mixtures with high dynamic range. Thus, high-performance liquid chromatography (HPLC) coupled with high-resolution mass spectrometry (HRMS) analytical method is developed and used in this dissertation for analysis lignin related compounds.

1.4 High-Performance Liquid Chromatography

High-performance Liquid chromatography (HPLC) is used in this dissertation as the main analytical separation technique. It is optimum for the separation variety of compounds that are non-volatile. HPLC is a type of liquid chromatography that provides good separation performance and it is run under high pressure. A typical HPLC system consists of a pump, an injector, and a separation column, reservoirs for liquid mobile phase and a detector. A separation column is usually packed with the stationary phase, which are very small porous particles (1-5 µm) with chemically modified surface resulting in the types of LC separation. Different chemical modification on the stationary phase generates a wide variety of LC separation techniques that are suitable for different analyte of interests. Among these, reversed-phase chromatography is widely used for many small organic molecules as well as peptides when using mass spectrometry as detector. Reserved-phase HPLC (RP-HPLC) is the most common HPLC technique because a large number of compounds can be separated with RP-HPLC. The chromatographic separation is performed on a non-
polar stationary phase with a polar mobile phase. There is a wide range of non-polar stationary phase available, which provides various options for choosing the correct separation system according to the physical and chemical properties of the compounds in order to achieve good separation. The stationary phase for RP-HPLC is usually obtained through chemically modified on the solid surface, such as silica. The most common one is C18 column, which is putting long chain hydrocarbon containing eighteen carbons giving very high hydrophobicity. Modification of solid surface with phenyl group is often used in order to separate aromatic compounds. The mobile phase used in RP-HPLC is usually a mixture of an organic solvent (acetonitrile, methanol, etc.) and water. A gradient of mobile phase is usually used when separation of complex compound mixture is needed. During RP-HPLC, a small volume of liquid sample is injected and loaded on the top of the column. The sample components then interact with the stationary phase through hydrophobic interaction. The process of the separation of the sample each component is determined by the choice of the chemical property of the column packing as well as the mobile phases

1.5 Electrospray Ionization- Orbitrap Mass Analyzer

The applications of mass spectrometry on lignin biopolymers are mainly focused on gas chromatogram (GC)/mass spectrometry coupled with chemical degradation or pyrolysis. It has drawn a lot of interest in developing MS-based techniques that would extend the research on studying the intact lignin polymers. Along with the development of soft-ionization techniques, such as electrospray ionization (ESI) and Matrix Laser Desorption Assisted Ionization (MALDI), application of mass spectrometry has been
extended to various compounds with wide range of molecular weight and provided useful structural information. In this dissertation, ESI-MS is the major analytical technique used to study and characterize the compounds of interest.

**Electrospray Ionization**

Electrospray ionization (ESI) is one of most extensively used ionization methods in mass spectrometry (MS). ESI has been mostly coupled with high-performance liquid chromatograph (HPLC) for the following reasons: a) the relative simplicity of ESI in introducing ions from solution phase to gas phase; b) wide range of eluent flow rate from a few µl/min to several hundred µl/min and c) applicability to a large number of analytes with diverse chemical and physical properties. Electrospray ionization takes place at atmospheric pressure. It is a “soft-ionization” method since there is little or no fragmentation during the electrospray ionization process. ESI can be used in both positive ion mode and negative ion mode. Here is the example of ESI process in positive ion mode. During ionization process (Figure 1.7), a solution of analyte is passed through a thin metal capillary which is applied several kV voltage (e.g., 500V to 4.5kV)\(^9\). The metal capillary is held positive potential vs. ground on the mass spectrometer, the high electric field at the emitter tip drives positively charged analytes to move toward the tip of the capillary emitter, resulting in an increasing density of positive ions at the liquid/air interface at the capillary tip. Under the effects of the applied electric field, once the coulomb repulsion between positive ions exceeds the surface tension of the liquid, a jet of positively charged solution emerges from the emitter tip, known as a Taylor cone\(^10\). As a result of formed Taylor cone, charged droplets ejects from the emitter. The sizes of these
droplets are around mm diameter range. The solvent then continuously evaporates off from the droplets owing to the high surface area-to-volume ratio, assisted by nebulizer gas flow and super-heated interface. The evaporation of solvent from initial charged droplets results in the shrinking of the droplets. As the droplets become smaller and smaller, the charge-to-volume ratio increases until the resulting coulomb repulsion overcomes the surface tension of the droplet, so called Rayleigh limit. The droplet reaching the Rayleigh limit forms a jet of offspring droplets, which are smaller and highly charged. The offspring droplets then undergo numbers of evaporation/jet fission cycles until a radius of about 5-10 nm is reached. Eventually gas-phase ions are formed\textsuperscript{10-11}. Several important parameters, such as spray voltage, flow rate, emitter tapering, surface hydrophobicity and fluid conductivity, play important roles on ionization efficiency of analytes.
Figure 1.7 Schematic depiction of an ESI source operated in the positive ion mode.
**Orbitrap mass analyzer**

Mass spectrometry is quite unique among other analytical instrumental methods, as several different physical principles of ions form the foundation for various types of mass analyzers. For instance, mass-to-charge ratio can be measured based on trajectories in a magnetic sector mass analyzer, path stability in a quadrupole mass filter mass analyzer, orbital frequency in an ion cyclotron resonance mass analyzer and velocity in a time-of-flight mass analyzer. Each mass analyzer has its own advantages and disadvantages.

The Orbitrap mass analyzer, used extensively in this dissertation, was first commercialized in 2005 by Thermo Fisher Scientific. It is a relatively new mass analyzer compared with other types, however, it quickly dominates the market because it can provide much higher mass accuracy and mass resolution compared to linear ion trap and is at a much lower cost than Fourier Transform Ion Cyclotron Resonance (FT-ICR).

In 1923, Kingdon first built an ion-trapping device using only electrostatic field based on a straight wire along the axis of outer cylindrical electrode. Later Knight refined the shapes of outer electrode. This trap design is often called “ideal Kingdon trap”. Makarov invented and filed a patent of the new type of mass analyzer, the Orbitrap, in 1999.

The ion-trapping in Orbitrap also uses electrostatic field but with specially shaped inner and outer electrodes compared with Kingdon trap. It has a spindle-like central electrode and a barrel-like outer electrode (Figure 1.8). The ions trapped in the Orbitrap can be described with potential distribution:
\[ U(r, z) = \frac{k}{2} \left( z^2 - \frac{r^2}{2} \right) + \frac{k}{2} (R_m)^2 \ln \frac{r}{R_m} + C \]  

(1.1)

Where \( r \) and \( z \) are cylindrical coordinates, \( C \) is a constant, \( k \) is field curvature and \( R_m \) is the characteristic radius. This field is quadrologarithmic because it is a sum of a quadrupole field of the ion trap and a logarithmic field of the cylindrical capacitor\(^\text{13}\).

In the field (1.1), ions move in an intricate spiral trajectory. By solving the equation of motion in polar coordinates \((r, \varphi, z)\) for ions with mass-to-charge ratio \( m/q \), three characteristic frequencies of the ion can be derived:

\[ \omega = \sqrt{\left( \frac{q}{m} \right) k}, \]  

(1.2)

where \( \omega \) is the frequency of axial oscillation\(^\text{13}\)

\[ \omega_r = \omega \sqrt{\frac{(R_m)^2 - 1}{2}}, \]  

(1.3)

where \( \omega_r \) is the frequency of radial oscillations\(^\text{13}\)

\[ \omega_\varphi = \omega \sqrt{\frac{(R_m)^2 - 1}{2}} \]  

(1.4)

where \( \omega_\varphi \) is the frequency of rotation\(^\text{13}\).

Out of the three characteristic frequencies, only the frequency of the harmonic axial oscillation \( \omega \), which is inversely proportional to the square root of the mass-to-charge ratio \( m/q \) of the ions, is completely independent of the energy and the position of ions. Therefore, only this frequency is used for mass analysis.
Figure 1.8. Simplified Schematics of an Orbitrap
Similar with FT-ICR instrument, image current detection is used as detector in Orbitrap. Image current detection relies in that an ion cloud attracts (positive ions) or repels (negative ions) the electrons of the detection electrodes when it passes split outer electrodes. The image current induced by the oscillation of ions can then be amplified, transformed into a voltage signal and translated into frequency domain by Fourier Transform resulting in accurate reading of ions’ mass-to-charge ratios. Because frequency can be measured at the highest accuracy among other physical quantities, Orbitrap can offer very high mass measurement accuracy. With the use of an internal calibration ion, mass accuracy can be achieved with 3 ppm. As the harmonic axial oscillation frequency of each m/z is translated from image current, the more transients are acquired in each scan, the higher resolving power of mass analyzer can be achieved. Thus, the time of analysis increases proportionately with the resolution required. Not only the analysis time is critical for high resolving power in Orbitrap, but it also requires ultra-high vacuum (<2x10^{-10} mbar) in order to allow ions to travel sufficiently long transients. The ultra-high vacuum is needed to provide the sufficient mean free path of hundreds of ions revolution undisturbed around the center electrode. However, ion collides with background gas molecules even at this ultra-high vacuum level, resulting in loss of phase coherence of ion packet, causing a decrease in the signal intensity of the transient\textsuperscript{9,12}.

The Orbitrap mass spectrometer used in this dissertation is Thermo Scientific Q-Exactive Orbitrap. The schematic of the mass spectrometer is showed in Figure 1.9). Q-Exactive Orbitrap consists with several components: ESI ion source, S-lens, transfer octopole, quadrupole mass filter, a curve RF-only quadrupole known as C-trap, Orbitrap mass analyzer and high-energy collision cell (HCD).
generated at the ion sources need to pass through a series of components before analyzed in the Orbitrap. C-trap, which is a bent RF-only quadrupole, is used to accumulate, store and thermalize ions prior to the injection to Orbitrap. It is important component for the high performance of Orbitrap. Ions get collisional damping in the C-trap resulting in efficiently capturing and cooling of the ions. Accumulated ions then can be injected from C-trap to Orbitrap with narrow energy distribution and injection angle$^9,^{14}$. The Q-Exactive can acquire MS data at several different mass resolutions according to the specific analytical requirements. The higher the mass resolution, the lower the MS scan speed, vice versa. As mentioned previously, frequency is the physical property that can be measured at the highest accuracy. With the use of internal calibration ion, the mass accuracy is usually within 2 ppm. The elemental composition information can be achieved with high-resolution accurate mass data with the corresponding mass error. This is extremely useful in providing another level of information regarding identifying unknown compounds. In this dissertation, the high-resolution accurate mass data obtained has been greatly aided in eliminating the unrelated ions and structural identification and confirmation.
Figure 1.9 Schematic of Q-Exactive mass spectrometer (Thermo Scientific) (source: http://planetorbitrap.com/q-exactive#tab:schematic)
2.1 Introduction

2.1.1 Formation and Structure of Lignin

As mentioned in the previous chapter, lignin is a complex heterogeneous biopolymer. It is mainly composed of three phenylpropanoid monolignols: 4-hydroxy cinnamyl alcohol, coniferyl alcohol and sinapyl alcohol (shown in Figure 2.1). When incorporated into lignin, they are referred as hydroxyphenyl (H), guaiacyl (G) and syringyl (S). Conventional numbering on the ring (C1-C6) and Greek letters on the side chain (α, β, γ) are used to describe the different chemical linkage types. The amount of three monolignols varies in different species. The dominant monolignol in gymnosperm lignins is G-unit with minor amount of H-unit, whereas angiosperm dicot lignins contain both G- and S- units. H-units are mainly present in softwood lignin and slightly higher in grasses\textsuperscript{15}. These units are linked through carbon-oxygen ether and carbon-carbon linkages in various bonding types. During lignin biosynthesis, peroxidases initiate dehydrogenation of monolignols at the 4-OH position. The resulting phenoxy radical is resonance stabilized by delocalizing to O-4, C-1, C-3, C-5 and C-β positions. Resonance structures of the phenoxy radicals for coniferyl alcohol are depicted in Figure 2.2. Two phenoxy radicals are involved in dimer formation via radical coupling reaction between active radical positions. The C-β appears to be the most reactive site as the most abundant linkages in lignin involve in C-β position (e.g. β-O-4, β-β and β-5)\textsuperscript{7} (Figure 2.3). An
example of the oxidative coupling of coniferyl alcohol to form the most abundant β-O-4 bond is shown in Figure 2.4. This linkage involves a two-step reaction: 1) radicals at O-4 and C-β position couple and form a Cβ-O4 ether, 2) the quinone methide intermediate generated from the coupling is quite reactive and can easily accept additions of another nucleophiles to the C1=Cα double bond. In this case, one water molecule acts as the nucleophile, which introduces an aliphatic alcohol at Cα position. Because the random and nonspecific chemistry of radical coupling reaction, biosynthesized lignin polymers are very complex with unknown combination of all different possible linkage as well as different frequencies of monolignols.

2.1.2 Dehydrogenation Polymers (DHPs)

Klason first pointed out that lignin is a derivative from the oxidation of coniferyl alcohol, even though he didn’t have direct experimental data to support his view in 1908. Freudenberg and his coworkers then established the fundamental understanding of lignin chemistry. Freudenberg and his coworkers first successfully synthesized lignin-like dehydrogenation polymers (DHPs) of coniferyl alcohol in the presence of laccase and hydrogen peroxide. DHPs have been widely used as a good model system for researchers to understand the chemistry of lignin for several reasons: 1) compared with native lignin, DHP does not have interference from other plant components, such as carbohydrates; 2) by controlling the polymerization condition, one can alter the DHP product profiles which provide deeper understanding of the biosynthesis of natural lignin;
Figure 2.1. Structures of three monolignols: a) coniferyl alcohol(G); b) \( p \)-coumaryl alcohol(H); c) sinapyl alcohol(S)
Figure 2.2. Resonance structure of phenoxy radical of lignin monomer.
Figure 2.3. Three Examples of common bonding types presented in lignin; a) β-aryl ether (β-O-4); b) resinol (β-β); c) phenylcoumaran (β-5).
Figure 2.4. Example of β-O-4 linkage formation via radical coupling
3) by isolating individual oligomers, one can obtain well-characterized lignin model compounds to be used for further studies.

The literature teaches that the product profiles of DHP, including frequency of each linkage type, molecular distributions and etc., heavily depend on the reaction conditions. There are generally two different reaction modes: 1) the reactants are added in a slow and continuous manner, which is the so called Zutropf (ZT) method; 2) The Zulauf (ZL) method has all the reactants mixed at the same time. The reaction solvent, pH, temperature and the presence of polymerization templates also have large impacts on the molecular weight distribution and primary structures of DHPs. The slow and continuous way of adding monolignols and hydrogen peroxide is considered to be more similar with the process occurring in the plant. ZT method results in a higher content of β-O-4 linkages than ZL method\textsuperscript{18-19}. The \textit{in vitro} polymerization is done in an aqueous buffer system, such as phosphate buffer with pH 6.5. As the polymerization proceeds, the resulting oligomers precipitate from aqueous buffer solution because of low solubility of oligomers, creating an inhomogeneous reaction system. The precipitation of the oligomers limits the production of high degrees of polymerization. De Angelis and coworkers developed a polymerization reaction in the presence of cationic surfactant in phosphate buffer. The surfactant keeps the oligomers in the solution so that a higher degree of polymerization can be achieved\textsuperscript{20-21}. The structures of DHPs synthesized with or without surfactant are distinctly different from each other.
2.1.3 Methods for Characterization of Lignin

Due to the heterogeneous nature of lignin polymers, extensive efforts have been investigated in order to characterize lignin. Early analytical techniques for characterization of lignins usually involved in the isolation and degradation of lignin into small pieces and then the polymer structure was deduced from characterization of the pieces. Chemical degradation methods were the only way to obtain structural information before the advent of various sophisticated analytical instrumental methods, which have greatly aided the knowledge of lignin structure.

Chemical degradation reactions are used in determination of the monomeric composition of lignin, both qualitatively and quantitatively. Table 2.1 gives a review of different chemical degradation methods for characterization of lignin.

**Spectroscopic methods in lignin study**

There is no singular method to characterize lignins. Lignins have been extensively investigated with a variety of spectroscopic techniques: 1) Ultraviolet (UV) - Visible spectroscopic spectra provide adequately understanding of the substituent effect of benzene ring in lignin; 2) Vibrational spectroscopic methods, including infrared (IR), near infrared (NIR) and Raman spectroscopy, have been important in analysis of the functional groups in lignin; 3) Nuclear Magnetic Resonance (NMR) spectroscopy has been extensively applied to complex lignin polymers in providing structural information. The two-dimensional (2D) NMR techniques are very powerful in determining the relative abundance of each monomeric units and relative frequencies of each types of bonding linkage, but does not provide information regarding the specific sequence of lignin.
### Table 2.1 A brief review of different chemical degradation methods for structural characterization of native lignin

<table>
<thead>
<tr>
<th>Degradation Condition</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oxidative Degradation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alkaline Nitrobenzene Oxidation (NBO)</td>
<td>Lignins are oxidized by nitrobenzene in 2M NaOH and high temperature (160-180°C)</td>
<td>First used to confirm the aromatic nature of lignins by Freudenberg; simple reaction mixture; yield satisfactory amount of benzaldehyde of H, G and S units; High reproducibility of S/G ratios</td>
</tr>
<tr>
<td>Permanganate Oxidation</td>
<td>Initial peralkylation of phenolic hydroxyl groups with diethylsulfate or</td>
<td>Evaluation of free phenolic groups in lignin</td>
</tr>
<tr>
<td>Method</td>
<td>Chemistry</td>
<td>Advantages</td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Thioacidolysis</td>
<td>dimethylsulfate, followed by oxidation with permanganate and hydrogen peroxide</td>
<td>Simple; informative; high yield of monomer; routine and robust; information of unusual units</td>
</tr>
<tr>
<td></td>
<td>Thioacidolysis: Acid-catalyzed reaction using ethanethiol and boron trifluoride etherate results in β-O-4 cleavage</td>
<td></td>
</tr>
<tr>
<td>Derivatization Followed by Reductive Cleavage (DFRC)</td>
<td>Bromination of the benzylic position and acetylation of free hydroxyl group, followed by reductive cleavage of β-O-4 linkage with zinc dust,</td>
<td>Mild reaction condition; routine and simplicity; Preserve the stereochemistry of C-β;</td>
</tr>
<tr>
<td>Method</td>
<td>Description</td>
<td>Additional Information</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Ozonation</td>
<td>This method preserves the side chains and oxidizes aromatic rings to carboxylic groups</td>
<td>Exclusively target on lignin side chains; provide insight on the stereo structure of lignin side chains;</td>
</tr>
<tr>
<td></td>
<td>Acetylation of final products for GC analysis</td>
<td>Provide new structural data of lignin chemistry, such as γ-ester groups remain unmodified during degradation method</td>
</tr>
</tbody>
</table>
primary structure; 4) Mass Spectrometry (MS) is extensively used when coupled with gas-chromatography (GC) by analysis of chemical degradation products of lignin for monomeric composition information. This method has limited applications on small volatile degradation products.

Along with the development of soft ionization techniques coupled with increasingly sophisticated MS hardware, mass spectrometry shows promising application in extending the studies on high molecular weight lignin compounds, which cannot be analyzed using GC/MS. Moreover, structural elucidation of lignin compounds is possible using tandem mass spectrometry (MS/MS).

**Ionization of Lignin compounds**

Traditionally, mass spectrometry (MS) was most commonly coupled with gas chromatography (GC) to identify volatile components in lignin degradation products using electron ionization (EI) technique. However, for the lignin samples with higher molecular weights and high polarities, alternative ionization and separation methods were required to characterize those complex lignin compounds.

In spite of the increasing popularity of the soft-ionization mass spectrometry methods for a variety of synthetic and biological molecules, the analysis of lignin products using different soft ionization techniques is under-developed by comparison. Matrix-assisted laser desorption ionization (MALDI) is also a “soft-ionization” technique coupled with time-of-flight (TOF) mass analyzer. MALDI-TOF instrument is widely used in the study of large biopolymers for molecular weight (MW) determination. Several groups have published articles using MALDI-MS to determine the MW.
distribution of the natural extracted lignin and synthetic lignin model compounds and to monitor the degrees of enzymatic polymerization during *in vitro* synthesis of lignin-like polymers\textsuperscript{20,23-26}. However, the quality and reproducibility of MALDI-MS data are largely associated with sample preparation, the choice of MALDI matrix, and interferences from matrix ions, which has led to poor inter-laboratory consistency\textsuperscript{27}.

The application of mass spectrometry in lignin chemistry has been greatly expanded by direct analysis in solution using electrospray ionization, which minimizes the inhomogeneity in sample preparation and interferences between analyte and matrix observed in MALDI-MS.

De Angelis *et al.* applied ESI-MS to characterize an *in vitro* synthesized model of dehydrogenation polymers (DHPs) in 1999\textsuperscript{20}. They successfully observed up to octamer as ammonium adduct ion in positive ion mode. Each oligomer showed definite mass and high regularity of the oligomers, differing by 178 and 180 mass units, which indicated that there were two different reaction mechanisms involved in the polymerization, one was radical coupling and the other was nucleophilic addition of monolignol. Evtuguin *et al.* used ESI to characterize extracted lignin and commercial lignin monomeric and dimeric model compounds in the same year\textsuperscript{28}. Evtuguin *et al.* acquired ESI mass spectrum in the negative ion mode in the presence of ammonium hydroxide to aid the deprotonation during ESI. The lignin monomeric and dimeric models used in those studies all contained free phenoxy or carboxyl group, which gave good response in the negative ion mode with simple deprotonation. Due to the high complexity of natural lignin sample, the corresponding ESI mass spectra simply showed as a molecular weight distribution curve instead of discrete peaks. The clusters of ions with 164 - 197 mass unit
difference were observed in the mass range from 500 to 1500, which was consistent with average molecular weight of monomeric units\textsuperscript{28}. It was obvious that negative ion mode would be a preferable method due to the free, weak acidic phenoxy group in lignin, as the phenoxy alcohol can be easily deprotonated and provide [M-H]\(^-\) ion. Önnerud \textit{et al.}\textsuperscript{29} applied both MALDI and ESI-Fourier Transform (FT)- Ion Cyclotron Resonance (ICR) to study the thioacidolysis products of lignin model compounds and mill wood lignin (MWL) with both positive ion mode and negative ion mode. Interestingly, they have reported the observation of [M+H]\(^+\), [M+Na]\(^+\), and [M+K]\(^+\) ions of thioacidolysis products with and without the presence of free phenoxy group. It was clear that these dimeric compounds can easily form adduct ions with Na\(^+\) and K\(^+\). Morreel \textit{et al.} studied the fragmentation pattern of dimeric model compounds with characteristic linkage types in negative ion mode using ESI-MS\textsuperscript{30-31}. Their work shows the promising application in elucidation of primary lignin structures with negative ion mode ESI-MS.

As for electrospray ionization, the nature of the analyte (e.g. functional group, polarity and etc.) and the spray solution have huge impacts on the performance and ionization efficiency of the analytes of interest. However, little effort has been expended to investigate into the evaluation of different ESI conditions for lignin. Haupert \textit{et al.} published on the characterization of monomeric and dimeric lignin model compounds under different ionization conditions, including ESI and atmospheric pressure chemical ionization (APCI)\textsuperscript{29,32}. Even though the presence of the weak acidic phenoxy group would prefer the negative ion formation via deprotonation, in-source fragmentation can negatively affect the ionization response of model compounds in negative ion mode using traditional ESI ion source. Those model compounds formed sodiated adduct ions in the
positive ion mode with no effect of in-source fragmentation. However, the study\textsuperscript{32} also concluded that no structural information obtained by tandem mass spectrometry as the sodiated cations were very difficult to be fragmented by collision activated dissociation (CAD). They have suggested that negative ion mode ESI with NaOH as dopant gives better ionization response with no obvious in-source fragmentation. Dean \textit{et al.} published the study of the interaction of alkali metal ions with dimeric model compounds using optical spectroscopy coupled with mass spectrometry in 2015\textsuperscript{33}. This work provided insight regarding the configuration of binding site of lithium or sodium cations with dilignols each containing unique bonding type ($\beta$-O-4 and $\beta$-$\beta$)\textsuperscript{33}. DeBlase \textit{et al.}\textsuperscript{34} later applied the same approach to study the alkali cation interaction with tetralignol containing pure $\beta$-O-4 linkage. They reported that there was size effect of alkali cations on the three-dimensional structure of the tetralignol model compound\textsuperscript{34}.

It is common that ionization response during ESI process can be very different for different compounds as they have different structural properties. Although lignins are polymers of phenylpropanoid units sharing similar functionalities, one should be careful in examining and interpreting the mass spectrometric data as the presence of “random” linkages in lignin. The complexity and “randomness” of lignin polymers are likely to give rise to different response during MS analysis. Further understanding regarding the ionization behavior of lignin is needed in order to confirm the molecular weight study when using mass spectrometry. The goal of this chapter is to characterize the average ESI response of model trilignols consisting with different distinct bonding motifs. The hypothesis of the first project contains two parts: 1) the electrospray ionization response study of lignin model compounds with distinct chemical linkages will give better
understanding in the structural effect on the ionization behavior; 2) by comparative analysis of high-energy collision dissociation (HCD) tandem mass spectrometry (MS/MS) of trilignols in both positive and negative ion mode, linkage-specific fragmentation pattern maybe obtained, which can lead to useful structural information in “lignomic” sequencing.

2.2 Materials and Methods

2.2.1 Materials

Coniferyl aldehyde, horseradish peroxidase (Type VI, 250-330 units/mg solid), cetyltrimethylammonium bromide were obtained from Sigma Aldrich (St. Louis, MO). Sodium phosphate, monobasic and sodium phosphate, dibasic were obtained from Fisher Scientific (Pittsburgh, PA). Cetyltrimethylammonium sulfate \((\text{CTA})_2\text{SO}_4\) was synthesized according to Feitosa, E.\textsuperscript{35}. Gel permeation chromatography (GPC) was accomplished with two different columns; a preparative (1 cm x 100 cm (Kontes, Fisher Scientific, Pittsburgh, PA)) column packed with 90 cm Bio-Beads S-X1 (Bio-Rad, Hercules, CA) and a 300 mm x 7.5 mm PLgel analytical GPC column (Agilent, Santa Clara, CA). Isolated oligomers were analyzed on a PolymerX (100 X 4mm, 3um) HPLC column (Phenomenex, Torrance, CA). All the MS analyses were done on a Q-Exactive Orbitrap mass spectrometer (ThermoScientific, Waltham, MA). Nuclear magnetic resonance (NMR) spectra were acquired on Varian Model NOVA400 (400MHz) spectrometer or Varian Model NOVA600 (600MHz) if otherwise mentioned.
2.2.2 Methods

Synthesis of Coniferyl alcohol

Coniferyl alcohol was prepared by the reduction of coniferyl aldehyde\textsuperscript{36} using NaBH\textsubscript{4}. The product was purified by SiO\textsubscript{2} column to yield a pure light yellow solid with 85% yield. \textsuperscript{1}H NMR (400 MHz, Chloroform-\textit{d}) \(\delta\) 6.96 – 6.79 (m, 3H), 6.53 (d, \(J = 15.8\) Hz, 1H), 6.22 (dt, \(J = 15.8, 6.0\) Hz, 1H), 5.67 (s, 1H), 4.30 (dd, \(J = 6.0, 1.5\) Hz, 2H), 3.90 (s, 3H).

Synthesis of Dehydrogenation Polymers (DHPs)

Synthesis Condition 1: DHP in the presence of Cetyltrimethylammonium sulfate (DHP-CT)

Coniferyl alcohol (GOH) was dissolved in 27 mM cetyltrimethylammonium sulfate [[(CTA)\textsubscript{2}SO\textsubscript{4}]] (CTAS) solution (made in 10mM sodium phosphate buffer) to a final concentration of 5 mg/ml. Cetyltrimethylammonium sulfate was prepared according to procedure\textsuperscript{35,37}. Horseradish peroxidase (HRP type VI, 250-330 unit/mg) was added at a ratio of 1 mg HRP/1000 mg coniferyl alcohol. Then 10% hydrogen peroxide (1mg GOH/2 \(\mu\)l H\textsubscript{2}O\textsubscript{2} solution) was added and stirred for 5 min. The reaction was quenched by adding several drops of NaHSO\textsubscript{3} and extracted with ethyl acetate\textsuperscript{21}.

Synthesis Condition 2: DHP synthesis using Zutropf (ZT) method (DHP-ZT)

Synthesis of \(\beta\)-O-4 linkage enriched DHP followed reported methods\textsuperscript{18} with slight modification. In general, three different solutions were prepared: Solution A: 3 mg/ml coniferyl alcohol dissolved in acetone: 10 mM phosphate buffer (1:9), Solution B: 0.15%
wt of hydrogen peroxide in 10 mM phosphate buffer, with the same final volume of
solution A, Solution C: 1µg/ml of horseradish peroxidase (HRP type VI, 250-330
unit/mg, amount of HRP used: 0.5 µg HRP per 1mg starting coniferyl alcohol) solution in
10 mM sodium phosphate buffer (pH 6.5). Solution A and B were added at a rate of 50
ml/hour by a syringe pump to solution C, stirred at 0°C. The reaction was quenched after
5 min with 5% NaHSO₃. The precipitate that formed was centrifuged to remove the
supernatant. The precipitate was dried using speed-vac. The aqueous portion was then
extracted with EtOAc and brine solution.

**Isolation of trilignols**

Guaiacylglycerol-β, γ-bis-coniferyl ether [G(4-O-α)G(β-O-4)G] was isolated
from reaction condition 1. The DHP was separated by semi-preparative gel permeation
chromatography (GPC) column, packed with Bio-beads S-X1, 85 cm X 1 cm, THF was
used as the elution solvent. The flow rate was at 0.2 ml/min₃⁸. Fractions were collected
every 10 min. Because the large amount of the surfactant [(CTA)₂SO₄] (CTAS) used in
the reaction, the isolated trimeric fraction was further purified on silica gel to remove
residual surfactant. HRMS: [M+NH₄]⁺ = 556.2544, Elemental composition: C₃₀H₃₈O₉N,
calculated: C₃₀H₃₈O₉N, m/z 556.2541, Δppm=0.5149 ppm.

Guaiacylglycerol-β-dehydrodiconiferyl ether [G(β-O-4)G(β-5)G] and
guaiacylglycerol-pinoresinol ether [G(β-O-4)G(β-β)G] were both isolated as a mixture
from reaction done in condition 2 from GPC column, using the same condition as
described above. The two trilignols were further separated by silica gel, eluted with ethyl
composition: $C_{30}H_{38}O_{10}N$, calculated: $C_{30}H_{38}O_{10}N$, m/z 572.2490, $\Delta$ppm=0.6412 ppm; 
$G(\beta-O-4)G(\beta-\beta)G$: $[M+\text{NH}_4]^+ = 572.2493$, Elemental composition: $C_{30}H_{38}O_{10}N$,
calculated: $C_{30}H_{38}O_{10}N$, m/z 572.2490, $\Delta$ppm=0.4279 ppm

The structures of isolated trilignols (shown in Figure 2.5) were confirmed with NMR.
The trimers were dissolved in 500 $\mu$l acetone-d$_6$ for NMR analysis. Acetone solvent 
peaks used as internal reference peaks, $\delta$H/$\delta$C 2.05 (5)/29.62 (7), 206.68 (1).
Assignments for the spectra are according the database of lignin model compounds$^{39-40}$.
Structural identification was performed by 1D and 2D NMR experiment ($^1$H and HSQC).
The chemical shift of carbon was obtained from HSQC experiment. Table 2.2 shows the 
assignments of main $^{13}$C-$^1$H cross signals of three trilignols of corresponding $^1$H and 
HSQC spectra.

**High-performance liquid chromatography (HPLC)-MS analysis**

HPLC separation was carried out by using PolymerX-TM RP-1(100 X 4 mm, 3 mm) at the flow rate of 0.4 ml/min coupled to a Thermo Scientific Q-Exactive mass 
spectrometer with the electrospray ionization(ESI) in positive ion mode. Spray voltage: 3.0 kV, sheath gas: 20 arbitrary units (AU), capillary temperature: 280, full scan range: m/z 150-2000. All data were analyzed using Thermo Xcalibur.

**Direct infusion MS analysis**

All MS experiments were run on a Thermo Q-Exactive mass spectrometer equipped with ESI source. To standardize the MS analysis, all the solutions were made in 
$H_2O/ACN=1:1$ with 0.5 mM final salt concentration, with 0.1 $\mu$g/ml n-hexyl triphenyl
phosphonium bromide as internal standard in positive ion mode. For cation selectivity study, the sample was prepared in the solution of H$_2$O/ACN=1:1, with 0.5 mM mixed salts (NH$_4$Cl, LiCl, NaCl, KCl and RbCl). For the negative ion mode via deprotonation, the solution was made in H$_2$O/ACN=1:1 with 0.5 mM NH$_4$OH. For the negative ion mode via chlorine adduct, the solution was made in H$_2$O/ACN=1:1 with 0.5 mM NH$_4$Cl. The solution was directly infused into MS using ESI at a flow rate of 3 µl/min with a micro-syringe pump. ESI condition: a) negative ion mode: spray voltage: 3.1 kV, sheath gas: 2 arbitrary units (AU), capillary temperature: 250, full scan range: m/z 100-1500; b) positive ion mode: spray voltage: 3.8 kV, sheath gas: 2 AU, capillary, temperature: 200°C, full scan range: m/z 100-1500. All data were analyzed using Thermo Xcalibur.
Figure 2.5. Structures of three trimers: a) Guaiacylglycerol-β, γ-bis-coniferyl ether [G(4-O-α)G(β-O-4)G]; b) Guaiacylglycerol-β-dehydrodiconiferyl ether [G(β-O-4)G(β-5)G]; c) Guaiacylglycerol-pinoresinol ether [G(β-O-4)G(β-β)G]
Table 2.2. NMR chemical shift assignments of three trimers

<table>
<thead>
<tr>
<th></th>
<th>( \delta_C / \delta_H ) (ppm)</th>
<th>Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>G(4-O-(\alpha))G((\beta)-O-4)G</strong></td>
<td>80.86/5.48</td>
<td>( A\alpha )</td>
</tr>
<tr>
<td></td>
<td>85.10/4.57</td>
<td>( A\beta )</td>
</tr>
<tr>
<td></td>
<td>61.61/3.94</td>
<td>( A\gamma_1 )</td>
</tr>
<tr>
<td></td>
<td>61.61/3.85</td>
<td>( A\gamma_2 )</td>
</tr>
<tr>
<td></td>
<td>129.65/6.50</td>
<td>( B\alpha/C\alpha )</td>
</tr>
<tr>
<td></td>
<td>129.10/6.20</td>
<td>( B\beta/C\beta )</td>
</tr>
<tr>
<td><strong>G((\beta)-O-4)G((\beta)-5)G</strong></td>
<td>73.96/4.89</td>
<td>( A\alpha )</td>
</tr>
<tr>
<td></td>
<td>88.649/4.20</td>
<td>( A\beta )</td>
</tr>
<tr>
<td></td>
<td>88.50/5.60</td>
<td>( B\alpha )</td>
</tr>
<tr>
<td></td>
<td>54.80/3.53</td>
<td>( B\beta )</td>
</tr>
<tr>
<td></td>
<td>NA/6.50</td>
<td>( C\alpha )</td>
</tr>
<tr>
<td></td>
<td>130.95/6.23</td>
<td>( C\beta )</td>
</tr>
<tr>
<td><strong>G((\beta)-O-4)G((\beta)-(\beta))G</strong></td>
<td>73.96/4.88</td>
<td>( A\alpha )</td>
</tr>
<tr>
<td></td>
<td>88.649/4.20</td>
<td>( A\beta )</td>
</tr>
<tr>
<td></td>
<td>88.65/4.67</td>
<td>( B\alpha )</td>
</tr>
<tr>
<td></td>
<td>55.62/3.09</td>
<td>( B\beta )</td>
</tr>
<tr>
<td></td>
<td>72.39/3.80</td>
<td>( B\gamma_1 )</td>
</tr>
<tr>
<td></td>
<td>72.39/4.19</td>
<td>( B\gamma_2 )</td>
</tr>
</tbody>
</table>
2.3 Result and Discussion

2.3.1 Characterization of Two Different DHPs

As discussed above that the structural profiles of *in vitro* synthesized DHPs are highly dependent on reaction conditions. From comparison of DHPs by ZL and ZT synthetic conditions, both DHPs contain similar bonding types but differ in relative abundance of each linkage types. However, DHPs synthesized in the presence of CTAS (DHP-CT) gives rise to distinctly different linkage types and corresponding frequencies of the linkages compared to ZL and ZT methods. Figure 2.6 shows the mass spectra of DHP-CT and DHP-ZT in the positive ion mode. Each oligomer forms ammonium adduct \([\text{M+NH}_4]^+\) species during the direct infusion in the presence of ammonium acetate. Figure 2.6.a shows that DHP-CT displayed a highly regular oligomeric mass pattern, which differs by 178 and 180 amu alternatively. The increment of 178 amu indicates two dehydrogenated radical coupling. Increment of 180 amu indicates that ionic reactions occurred by the addition of monomer to a quinone methide intermediate. This leads to the formation of \(\alpha, \beta\)-diaryl linkage in the lignin. In contrast to DHP-CT, DHP-ZT contains higher percentage of \(\beta\)-O-4 linkage. Figure 2.6.b is the mass spectra of DHP-ZT at the same ionization condition with that of DHP-CT, it appears to be a more complex product profile compared with DHP-CT. The cluster-like mass spectra in DHP-ZT results from oligomers that have same number degree of polymerization (DPn) but contain different numbers of \(\beta\)-O-4 linkage.

DHP-CT and -ZT samples were first analyzed in the negative ion mode ESI with simple deprotonation, shown in Figure 2.7. Interestingly, dramatic difference of the
Figure 2.6. Full scan mass spectrum of DHP in Positive ion mode ESI, [M+NH₄]⁺; a) DHP-CT; b) DHP-ZT
Figure 2.7. Full scan mass spectrum of DHP in negative ion mode ESI, [M-H]; a) DHP-CT; b) DHP-ZT
molecular distribution of DHP-CT sample was observed compared with that in the positive ion mode. For DHP-ZT samples, the molecular distribution was similar in positive and negative ion mode, but an overall higher response in positive ion mode than that in negative ion mode was observed.

These observations prompted the question of which ionization conditions were better represented the actual composition of the sample mixtures. Size exclusion chromatography (SEC), also known as gel permeation chromatography (GPC), is often used for analysis of molecular distribution of lignin. Figures 2.8 shows the GPC chromatograms of separated oligomer components from coniferyl alcohol standard, dimer, trimer, tetramer and pentamer. The separation efficiency of GPC column was not high enough to have baseline separation of each oligomer. The peak at rt =5.85 min in the top chromatogram in Figure 2.8 is a polystyrene standard, which has a MW around 1500. As shown in the GPC chromatogram of DHP-CT after 10min reaction (Figure 2.9 top), the most abundant oligomer is trimer, there are some dimer and monomer, and for the higher DP oligomers could not achieve baseline separation with the GPC column used in this experiment. DHP-ZT, obtained from a 5 min reaction, shows similar molecular weight distribution with DHP-CT (Figure 2.9). GPC of DHP-ZT shows similar molecular weight distribution as observed in both positive ion mode and negative ion mode in ESI-MS. Positive ion mode ESI-MS of DHP-CT shows closer pattern with GPC result, while negative ion mass spectrum is significantly under-presented the molecular weight profiles of DHPs. This important observation motivated this study into what difference between two DHPs gave rise to the difference in ESI response. In order to quantitatively study the ionization response, trilignols from DHPs were isolated and used to specifically compare
the ionization response with the presence of unique linkage types. Because trilignols
contains more degrees of freedom than that of monomer and dimer, ionization behavior
differences could be more significant than that of dimers. Moreover, trimers were
separable using GPC and more abundant than other higher oligomers, because of the
specialty of the GPC column used in this study. It is worth to point out that the each of
the three isolated trilignols contains a mixture of diastereomers, which would be a good
representation for natural lignin, as the goal of this study is aimed to characterization the
average response of DHP with electrospray ionization.
Figure 2.8. GPC chromatogram of isolated DHPs; from top to bottom were coniferyl alcohol, dimers, trimers, tetramers and pentamers.
Figure 2.9. GPC chromatogram of DHPs from two different reaction conditions: DHP-CT products (top) and DHP-ZT products (bottom)
2.3.2 Negative Ionization Behavior of Trilignols

**Simple deprotonation [M-H]**

As shown in Figure 2.10.a, the absolute ion intensity of G(4-O-α)G(β-O-4)G is significantly lower than other two trilignols at the same concentration, while G(β-O-4)G(β-β)G gives the highest ionization response. Mass spectrum of G(4-O-α)G(β-O-4)G shows high abundance of ion 357.1341 and 179.0712 (m/z consistent with the presence of dimer and monomer, respectively), while a low ion signal of the target trimeric ion is observed at m/z 537.2130 (Figure 2.10.b). Ions of 179.0712 and 357.1341 are likely resulting from in-source fragmentation during the ionization process. In order to support the hypothesis of in-source fragmentation of G(4-O-α)G(β-O-4)G, an LC-MS experiment was conducted on G(4-O-α)G(β-O-4)G. The EIC chromatograms (Figure 2.11) shows that m/z 357.1341 and m/z 179.0712 co-elute with target trimer m/z 537.2130 and have identical chromatographic peak shape. Analysis of a standard of monomer coniferyl alcohol shows that coniferyl alcohol has an earlier retention time than that observed in the LC run of G(4-O-α)G(β-O-4)G. The comparison result of LC runs confirms that m/z 179.0712 observed in the spectrum of G(4-O-α)G(β-O-4)G is not an actual coniferyl alcohol but most likely from a fragment ion. Unlike G(4-O-α)G(β-O-4)G trimer, mass spectra of G(β-O-4)G(β-β)G and G(β-O-4)G(β-5)G show a very low amount of dimer ion, which could be because of the higher stability of cyclic structural moiety existing in these two trimers. Therefore, deprotonated G(4-O-α)G(β-O-4)G must undergo a charge-driven in-source fragmentation pathway because the phenolic monomeric moiety
serving as a very good leaving group compared to α-alcoholic group in typical β-O-4 group.

A proposed fragmentation pathway of G(4-O-α)G(β-O-4)G in negative ion mode is shown in Figure 2.12. The lone pair of electron resonated to α-C and the moiety C, shown in Figure 2.5.a), results in labile 4-O-α linkage during ionization. Even though for other two trilignols, it can also undergo this charge-driven fragmentation, the alcoholic group at α-C is much poorer leaving group compared to phenolic group at α-C in G(4-O-α)G(β-O-4)G. As a result, the in-source fragmentation has less effect on G(β-O-4)G(β-5)G and G(β-O-4)G(β-5)G than that on G(4-O-α)G(β-O-4)G.

**Chlorine ion adduct formation [M+Cl]⁻**

To test the hypothesis of in-source fragmentation, MS spectrum of G(4-O-α)G(β-O-4)G in the presence of NH₄Cl was acquired where [M+Cl]⁻ was the dominant ion and only a few dimeric ion and monomeric ion were observed. When forming chlorine ion adduct of G(4-O-α)G(β-O-4)G, unlike deprotonation, there is no free lone pair of electrons generated. As a result, the chlorine adduct anion does not undergo the proposed charge-driven fragmentation pathway. Chlorine adduct experiment supported the hypothesis regarding the activation of fragmentation via the deprotonation at the phenolic site.

As the preliminary data show that negative ionization by forming Cl⁻ adduct greatly decreased the extent of in source fragmentation of G(4-O-α)G(β-O-4)G. It is interesting to see whether three trimers will behave differently with formation of Cl⁻ adduct by comparing the Cl⁻ adduct of all three trilignols
Figure 2.10. a) Deprotonated ion intensity comparison of three trimers; b) full scan mass spectrum of G(4-O-α)G(β-O-4)G. Error bars indicate ± one standard deviation.
Figure 2.11. Liquid chromatograms of G(4-O-α)G(β-O-4)G, a) BPC of trimer 537.213; b) EIC of 537.213; c) EIC of dimer, m/z 357.134; d) EIC of monomer, m/z 179.071; e) EIC of coniferyl alcohol, 179.071
Unlike with deprotonation, all trilignols have similar ionization response by forming \( \text{Cl}^- \) adduct, shown in Figure 2.13. \([\text{M-H}]^-\) of G(\(\beta\)-O-4)G(\(\beta\)-5)G and G(\(\beta\)-O-4)G(\(\beta\)-\(\beta\))G are still observed in Cl\(^-\) solution, which result from in-source fragmentation ions. Compared to previous deprotonation ionization, ion behavior of Cl\(^-\) adducts shows better ionization response across all three trimers. The HCD/MS/MS of \([\text{M+Cl}]^-\) ions contain essentially the same fragmentation pattern with that of deprotonated ions. The loss of HCl gives rise to the main fragment ion of \([\text{M-H}]^-\). As the increasing of collision energy, \([\text{M-H}]^-\) is further under fragmentation, which has the same fragmentation pathway as those from deprotonated precursor ions. Formation of Cl\(^-\) adducts gives overall better response than simple deprotonation in the negative ion mode. It also seems promising in facilitating the formation of gas-phase negative adduct ions, all three trimers result in good ionization response with some extent of in-source fragmentation shown in Figure 2.13.
Figure 2.12. Proposed fragmentation mechanism of deprotonated G(4-O-α)G(β-O-4)G ion.
Figure 2.13. Ion intensity comparison of three trimers with chlorine ion adduct, [M+Cl]⁻ in negative ion mode ESI. Error bars indicate ± one standard deviation.
2.3.3 Positive Ionization Behavior of Trilignols

Previous ionization experiment of both different DHPs gives pretty good ion response in positive ion mode. Haupert et al.\cite{32} has reported that sodium adduct give good response because of no in-source fragmentation. It is worthwhile to compare ionization and fragmentation behavior of trimers across a series cations (NH$_4^+$, Li$^+$, Na$^+$, K$^+$ and Rb$^+$).

(a) G(4-O-$\alpha$)G(\beta-O-4)G

Comparison of ionization in NH$_4$Cl and NH$_4$OAc was first done to see whether a counter ion would have effect on the ionization response. There is no significant difference in [M+NH$_4$]$^+$ formation in presence of two different counter anions Cl$^-$ and OAc$^-$ at p<0.05. The ion intensities of both lithiated and sodiated ions are higher than the other three cation adducts (shown in Figure 2.14.a). In the positive ion mode, very few dimeric and monomeric ions are observed, which is opposite to what is observed in the negative ion mode. It is likely that cationization increases the rigidity of G(4-O-$\alpha$)G(\beta-O-4)G trimer, preventing it from in-source fragmentation\cite{32}. Figure 2.14.b shows the ion competition studies which G(4-O-$\alpha$)G(\beta-O-4)G is dissolved in a mixture of different cations with equal molar concentration. The ionization competition experiment shows similar response pattern with the individual cation experiment, where Li$^+$ and Na$^+$ are preferred in the ion competition experiment.
Figure 2.14. a) Relative ionization response of G(4-O-α)G(β-O-4)G in different salt solution. b) Cation selectivity of G(4-O-α)G(β-O-4)G in solution of equal molar of NH₄Cl/LiCl/NaCl/KCl/RbCl. Error bars indicate ± one standard deviation.
(b) G(β-O-4)G(β-5)G

Figure 2.15a shows the ion response of G(β-O-4)G(β-5)G in different salt solutions. Again, there is no significant effect in counter ions when compared the ion response of ammonium adduct formation in the presence of NH₄Cl and NH₄OAc. No significant difference is observed in NH₄⁺, Li⁺, Na⁺ and K⁺ adducts, while Rb⁺ adduct formation has the lowest response among all cations.

Figure 2.15b shows the ion competition result of G(β-O-4)G(β-5)G exposed to the mixture of salts. The relative abundance of [M+Li]⁺, [M+Na]⁺, [M+K]⁺ and [M+Rb]⁺ is similar to the ionization comparison in individual solution (Figure 2.15.a). G(β-O-4)G(β-5)G shows most affinity to Li⁺, Na⁺ and K⁺. [M+Rb]⁺ formation is less favorable compared to the formation of [M+Li]⁺, [M+Na]⁺, [M+K]⁺. Unlike the individual ionization comparison, [M+NH₄]⁺ gives lowest ion abundance compared to the other four cation adducts.

(c) G(β-O-4)G(β-β)G

The ionization response of G(β-O-4)G(β-β)G in different salt solutions is shown in Figure 2.16.a. Again, no significant counter ion (OAc⁻ and Cl⁻) effect is observed by comparing ammonium adduct formation on G(β-O-4)G(β-β)G. Ion responses of lithiated and sodiated adduct ions show no difference. They both give higher response than NH₄⁺, K⁺ and Rb⁺. Among all cationization, NH₄⁺ resulted in the lowest ion response. Figure 2.16.b shows that the cation competition results follow similar trends with the individual ionization comparison. NH₄⁺ adduct formation is the least favored with G(β-O-4)G(β-β)G.
The data shows that different trimeric structures give different cationic response in the presence of different cations. Not only it is worthwhile to see whether individual trimer has a response preference toward particular cation, but it is more interesting to compare whether there is difference between three trilignols in the cationization response from each other due to their structural difference. Such comparison is shown in Figure 2.17. In contrast to the ionization response in the negative ion mode ESI ([M-H]), G(4-O-α)G(β-O-4)G is much more responsive compared to the other two trilignols in most of cations. In general, it shows that the relative ionization response of three trilignols for all the cations followe the order of G(4-O-α)G(β-O-4)G > G(β-O-4)G(β-β)G > G(β-O-4)G(β-5)G
Figure 2.15. a) Relative ionization response of G(β-O-4)G(β-5)G in different salt solution. b) Cation selectivity of G(β-O-4)G(β-5)G in solution of equal molar of NH₄Cl/LiCl/NaCl/KCl/RbCl. Error bars indicate ± one standard deviation.
Figure 2.16. a) Relative ionization response of G(β-O-4)G(β-β)G in different salt solution. b) Cation selectivity of G(β-O-4)G(β-β)G in solution of equal molar of NH₄Cl/LiCl/NaCl/KCl/RbCl. Error bars indicate ± one standard deviation.
Figure 2.17. Positive ion mode ESI response comparison of three trilignols with the presence of different cations. Error bars indicate ± one standard deviation.
2.3.4 HCD/MS/MS Experiment of Trilignols

Tandem mass spectrometry has been successfully applied in sequencing the primary structure of proteins, similar strategy to sequence lignin is proposed by using tandem mass spectrometry. Morreel et al. has done wonderful work in studying the signature fragment ions of the dimeric model compounds containing different linkage types in the negative ion mode\(^{30-31}\).

Previous experiments have shown that positive ion mode gives overall better response than negative ion mode. However, Haupert et al. has found it difficult to obtain fragmentation ions with sodiated adduct analyte with collision activated dissociation (CAD) in the tandem mass spectrometry\(^{32}\). Moreover, alkali metal adduct ions are considered to be problematic in tandem mass spectrometry previously using CID. No useful fragmentation ions can be obtained because the dissociation of alkali metal ion itself is preferred during collision induced dissociation (CID). Because the difference in the fundamental dissociation mechanism between CID and high-energy collision dissociation (HCD), different fragmentation of alkali metal adducts may be observed using HCD. It will be beneficial to study whether high-energy collision dissociation/tandem mass spectrometry (HCD/MS/MS) experiment will provide useful structural information. If so, one may then distinguish the linkages via the HCD/MS/MS spectra. The accurate MS data are acquired which provide additional confident in analyzing the MS/MS data. For all three trilignols, HCD/MS/MS spectra of \([\text{M+NH}_4]^+\) contain most complicate fragment ions than other alkaline metal ion adducts. \(\text{NH}_3\) is easily knocked off from ammonium adduct ion, the resulting fragment ions are protonated. Compared with HCD/MS/MS of other cation adducts, MS/MS of
Figure 2.18. HCD/MS/MS spectra of three trimers with ammonium adduct, [M+NH₄]^+, a) G(4-O-α)G(β-O-4)G; b) G(β-O-4)G(β-5)G; c) G(β-O-4)G(β-β)G.
[M+NH₄]⁺ are more complicated to study signature fragment pattern of different linkages, so the MS/MS spectra of [M+NH₄]⁺ for three trilignols will not be further discussed (shown in Figure 2.18).

**HCD/MS/MS spectra of G(4-O-α)G(β-O-4)G**

[M-H]⁻

As discussed previously, α, β-diaryl ether bond is quite labile compared with other C-C condensed linkages; it is easily fragmented to monomeric and dimeric fragment ions (shown in Figure 2.19.a.) at lower collision energy compared with other two trimeric compounds. The dominant fragmentation ions, m/z 357 and 179 observed in the HCD/MS/MS experiment are the same with the ones from in-source fragmentation of [M-H]⁻ ion. Besides the major dimeric (m/z 357) and monomeric (m/z 179) fragmentation ions, small neutral losses of 18 Da (H₂O), 30 Da (CH₂O) are also observed. The loss of water (18 Da) indicates the presence of aliphatic alcohol group. The loss of formaldehyde (30 Da) also indicates the existence of aliphatic alcohol group.

[M+Li]⁺ and [M+Na]⁺

Fragmentation of lithiated and sodiated ions are required higher collision energy than that of [M-H]⁻. As shown in Figure 2.19.b-c, HCD/MS/MS spectra of both lithiated and sodiated ions give essentially the same fragmentation ion. Both MS/MS spectra are pretty clean, containing only two major fragment ions, which correspond to the loss of a monomeric unit and the loss of a dimeric unit (loss of 180 and 359). The fragmentation patterns of
Figure 2.19. HCD/MS/MS spectra of G(4-O-α)G(β-O-4)G: a) [M-H]⁻; b) [M+Li]⁺; c) [M+Na]⁺.
lithiated and sodiated ions are very similar with that of deprotonated ions. However, there are no loss of 18 Da and 30 Da observed in these two HCD/MS/MS spectra.

**HCD/MS/MS spectra of G(β-O-4)G(β-5)G**

**[M-H]**

Compared with G(4-O-α)G(β-O-4)G, G(β-O-4)G(β-5)G is more stable when subjected to fragmentation. The tandem spectrum is also more complicated than that of G(4-O-α)G(β-O-4)G. At the same collision energy, shown in Figure 2.20.a, the base peak is [M-H], m/z 553.2074(C_{30}H_{33}O_{10}, Δppm= -1.06, 100%). m/z 195.0661(C_{10}H_{11}O_{4}, Δppm= -1.0846, 82.36%), which indicates the existing of β-O-4 at phenolic end group of guaiacyl unit. m/z 339.1232 (C_{20}H_{19}O_{5}, Δppm= -1.85, 8.27%), m/z 327.1234(C_{19}H_{19}O_{5}, Δppm= -1.17, 28.94%) are two characteristic fragment ions corresponding to the presence of G(β-5)G linkage\textsuperscript{30}. The common fragmentation ions with loss of 18 Da and 30 Da are also observed which indicates the presence of aliphatic alcohol groups as discussed previously.

**[M+Li]**\textsuperscript{+} and **[M+Na]**\textsuperscript{+}

Lithiated and sodiated of G(β-O-4)G(β-5)G generally are more difficult to be fragmented, thus they require higher collision energy than that of α, β-diaryl trimer. As showed in Table 2.3, fragmentation ions of lithiated and sodiated ion overlap with each other pretty well. Not only fragmentation ions of alkaline cation adduct ions are obtained, the characteristic fragmentation ions are also observed in the HCD/MS/MS of lithiated
Figure 2.20. HCD/MS/MS spectra of G(β-O-4)G(β-5)G: a) [M-H]; b) [M+Li]; c) [M+Na].
Table 2.3. List of observed fragment ions for G(β-O-4)G(β-5)G and their chemical formula based on high resolution MS result of [M+Li]$^+$ and [M+Na]$^+$

<table>
<thead>
<tr>
<th>Observed Ion</th>
<th>[X+Li]</th>
<th>[X+Na]</th>
</tr>
</thead>
<tbody>
<tr>
<td>C30 H34 O10</td>
<td>561.2308(100%)</td>
<td>577.2047(100%)</td>
</tr>
<tr>
<td>C30 H32 O9</td>
<td>543.2201(16.99%)</td>
<td>559.1942(6.68%)</td>
</tr>
<tr>
<td>C29 H32 O9</td>
<td>531.2205(14.29%)</td>
<td>547.1941(10.44%)</td>
</tr>
<tr>
<td>C29 H30 O8</td>
<td>513.2099(4.49%)</td>
<td>529.1836(1.78%)</td>
</tr>
<tr>
<td>C20 H21 O6</td>
<td>-</td>
<td>380.1231(10.14%)</td>
</tr>
<tr>
<td>C20 H19 O5</td>
<td>346.1388(3.56%)</td>
<td>362.1125(8.92%)</td>
</tr>
<tr>
<td>C19 H19 O5</td>
<td>-</td>
<td>350.1125(6.92%)</td>
</tr>
<tr>
<td>C17 H20 O6</td>
<td>327.1415(7.29%)</td>
<td>-</td>
</tr>
<tr>
<td>C17 H18 O5</td>
<td>309.1309(15.99%)</td>
<td>325.1047(33.94%)</td>
</tr>
<tr>
<td>C16 H15 O4</td>
<td>-</td>
<td>294.0863(8.92%)</td>
</tr>
<tr>
<td>C13 H16 O5</td>
<td>259.1153(6.92%)</td>
<td>275.089(3.59%)</td>
</tr>
<tr>
<td>C10 H12 O4</td>
<td>203.0891(2.06%)</td>
<td>-</td>
</tr>
<tr>
<td>C10 H11 O3</td>
<td>186.0864(1.93%)</td>
<td>202.0602(11.04%)</td>
</tr>
<tr>
<td>C8 H8 O3</td>
<td>-</td>
<td>175.0367(1.89%)</td>
</tr>
<tr>
<td>C8 H8 O2</td>
<td>143.068(4.31%)</td>
<td>-</td>
</tr>
</tbody>
</table>
and sodiated cations, which allowed identification and elucidation of detailed structure. Similar characteristic fragment ions are obtained in the negative ion mode for structure identification\textsuperscript{31}.

**HCD/MS/MS spectra of G(β-O-4)G(β-β)G**

\[ [M-H]^− \]

Morreel, K \textit{et al.}\textsuperscript{30} has studied the fragmentation pattern of deprotonated G(β-O-4)S(β-β)G, which it has a syringyl unit in the molecules. Again, the major fragment ions of deprotonated G(β-O-4)G(β-β)G follow the similar fragmentation pattern with their study. Because the HCD/MS/MS data are obtained in the HCD cell, there may have difference in the relative abundance of fragment ions compared with CID/MS/MS.

In Figure 2.21.a, which is the HCD/MS/MS spectrum of deprotonated ion \([M-H]^−\), the base peak is \([M-H]^−, m/z 553.2074(C_{30}H_{33}O_{16}, \Delta ppm= -0.9453, 100\%)\), 503.1867(C_{29}H_{29}O_{8}, \Delta ppm= -0.2277, 32.22\%), the signature fragment ions are 357.1339(C_{20}H_{21}O_{6}, \Delta ppm= -1.18, 11.33\%), 343.1182(C_{19}H_{19}O_{6}, \Delta ppm= -1.42, 0.26\%), 327.1234(C_{19}H_{18}O_{5}, \Delta ppm= -1.17, 0.92\%) and 195.066(C_{10}H_{11}O_{4}, \Delta ppm= -1.4, 29.07\%), which again indicates the existing of β-O-4 at phenolic end group.

\[ [M+Li]^+ \text{ and } [M+Na]^+ \]

Similar to G(β-O-4)G(β-5)G trimer, the fragmentation of lithiated and sodiated G(β-O-4)G(β-β)G ions also require higher collision energy than that of G(4-O-α)G(β-O-4)G. Table 2.4 shows the HCD/MS/MS comparison of \([M+Li]^+ \text{ and } [M+Na]^+ \) of G(β-O-4)G(β-β)G. Different with G(β-O-4)G(β-5)G trimer, HCD/MS/MS spectrum of
Figure 2.21. HCD/MS/MS spectra of G(β-O-4)G(β-β)G: a) [M-H]+; b) [M+Li]+; c) [M+Na]+.
Table 2.4. List of observed fragment ions for G(β-O-4)G(β-β) and their chemical formula based on high resolution MS result of [M+Li]$^+$ and [M+Na]$^+$

<table>
<thead>
<tr>
<th>Observed ions</th>
<th>[X+Li]</th>
<th>[X+Na]</th>
</tr>
</thead>
<tbody>
<tr>
<td>C30 H34 O10</td>
<td>561.2308 (100%)</td>
<td>577.2049 (100%)</td>
</tr>
<tr>
<td>C29 H30 O8</td>
<td>513.2098 (33.88%)</td>
<td></td>
</tr>
<tr>
<td>C20 H24 O7</td>
<td>383.1677 (6.44%)</td>
<td></td>
</tr>
<tr>
<td>C20 H22 O6</td>
<td>365.1572 (7.29%)</td>
<td></td>
</tr>
<tr>
<td>C20 H21 O6</td>
<td></td>
<td>380.1233 (10.33%)</td>
</tr>
<tr>
<td>C19 H18 O6</td>
<td>349.1258 (2.1)</td>
<td>365.0999 (2.65%)</td>
</tr>
<tr>
<td>C17 H18 O5</td>
<td>309.1308 (1.82%)</td>
<td></td>
</tr>
<tr>
<td>C12 H12 O3</td>
<td>221.0997 (3.31%)</td>
<td></td>
</tr>
<tr>
<td>C10 H12 O4</td>
<td>203.0891 (12.15%)</td>
<td></td>
</tr>
<tr>
<td>C10 H11 O3</td>
<td>186.0864 (6.02%)</td>
<td>202.0603 (11.5%)</td>
</tr>
</tbody>
</table>
lithiated ion contains more characteristic fragment ions than that of sodiated ion. Almost all the ions observed in MS/MS of [M+Na]^+ are also observed in [M+Li]^+, shown in Figure 2.21.b-c.

**Comparison of Fragmentation of three trilignols**

Fragmentation in HCD/MS/MS experiments takes much higher collision energy for alkali metal adducts than that of ammonium adduct ions in the positive ion mode. The HCD/MS/MS of [M+NH_4]^+ contains predominant [M+H]^+ fragmentation ions, as ammonium adduct is easily fragmented to H^+ and NH_3 leaves as neutral. Although alkali metal adducts require higher collision energy to obtain informative fragmentation ion, the HCD/MS/MS spectrum is relatively less complicated to interpret. Compared with HCD/MS/MS spectra of three trilignols, the lithiated and sodiated trimers have generated characteristic fragmentation ions and they match with the fragmentation pattern obtained in the negative ion mode (both [M-H]^- and [M+Cl]). The relative abundance of fragment ions is quite different from that of deprotonated ion. The HCD/MS/MS results of the trilignols cation adducts indicates that the fragmentation pathway varies with the species of cation as well as the structures of trilignols. By interpreting the fragmentation pattern of the lithiated and sodiated G(β-O-4)G(β-5)G and G(β-O-4)G(β-5)G adducts, one can also be able to differentiate these two isomers.

**2.4 Conclusion**

Because of the heterogeneity of lignin, an efficient ionization method for mass spectrometric analysis is needed to improve the sensitivity as well as obtaining structural
information of lignin. By studying trilignols as lignin model compounds, it has been demonstrated that how the characteristic structural has an effect on traditional negative ion mode ESI. In this study, it is shown that the ion response in the positive ion mode forming various cations adducts is overall more responsive than that in the negative mode. In general, Li$^+$ and Na$^+$ adduct ions give the highest ion response across three different trilignols, they can provide molecule weight information and be less effected by in-source fragmentation during the ionization process. Not only are [M+Li]$^+$ and [M+Na]$^+$ fragment ions successfully observed in HCD/MS/MS experiments, they also provide informative structural and sequence information for potential structural elucidation application. Cationization study of these model compounds containing phenolic alcohol may be extended for ionization of lignin oligomers whose phenol groups are occupied resulting difficulties in deprotonation. Application of cationization shows promise with good ionization response as well as structurally informative HCD/MS/MS data on lignin oligomers that contain various chemical linkages.
3.1 Introduction

3.1.1 Lignin model compounds

Lignin is the second most abundant biopolymer on the earth. As described in the previous chapter, lignin has very complex polymeric structure due to: 1) heterogeneous biopolymer consisting of three phenylpropanoid monomeric units: 4-hydroxy-cinnamyl alcohol, coniferyl alcohol and sinapyl alcohol; 2) biosynthesis through oxidative radical coupling reaction resulting in non-specific pattern of bonding types ($\beta$-O-4, $\beta$-5, $\beta$-1, $\beta$-$\beta$ and etc.); and 3) intra-molecular bonding with carbohydrates. For pulp/paper and cellulosic biofuel industries, cellulose is the component of interest. As an unwanted pretreatment byproduct, lignin is often burned for fuel. The complexity of the isolated lignin itself makes it very difficult to be commercialized as a highly valuable product. However, lignin has drawn a lot of attention because of the valuable aromaticity of lignin as a potential alternative resource of aromatic rings other than non-renewable fossil fuels. This potential for lignin to replace petroleum as a source of aromatics motivates a growing interest in understanding the fundamental chemistry of lignin degradation processes in order to obtain the highly valuable lignin-derived products.

As discussed previously, $\beta$-O-4 ether linkage is considered to be the most abundant linkage in lignin comprising from 40-60% of lignin depending on the type of plants. Compared with other C-C condensed linkage, $\beta$-O-4 is more labile, which makes it a reactive target site in lignin degradation in many degradation strategies.
Dimeric model compounds containing β-O-4 have been widely studied, but these cannot properly represent the polymeric property of lignin. Dehydrogenation polymer (DHP) as models have been discussed in Chapter two of this research program as a more “controllable” model but its structures are still too complicated to be used as a proper model to study lignin degradation process. Several groups have synthesized oligomeric model compounds containing β-O-4\(^{41-45}\). Some of these model compounds do not contain phenoxy functionality; some of them do not have the end group that maintains the unsaturated side chain. Thus, the objective of this chapter is to develop a synthetic method so that one can obtain oligomeric models that would contain important functionalities (phenolic alcohol end group, β-O-4 ether linkage as well as α, β-unsaturated side chain with γ–OH end group). Successful synthesis of a model compound with the previously mentioned features would have a significant impact on understanding the reactivity of the lignin model compounds towards degradation methods.

The strategy used to generate β-O-4 ether linkage is through an aldol-type reaction\(^4\). Ciofi-Baffoni et al. used lithium diisopropylamide as strong base to generate a carbanion as nucleophile addition to the aldehyde group. However, this condition will not work if one wants to introduce α, β-unsaturated side chain to the oligomeric models. The aim of this work is to develop the new synthesis strategy to prepare β-O-4 oligomeric model compounds that contain the most important functional groups: phenolic alcohol, β-O-4 and unsaturated side chain, which would be a good model for studying the reactivity and chemistry of lignin degradation process. The reaction synthesis route designed in this chapter is shown in Figure 3.1.
Figure 3.1. Synthetic route of target trimer G(β-O-4)G(β-O-4)G
3.1.2 Condensation Reaction of Aldehydes and Ketones

The idea of generating β-O-4 linkage is achieved by the reduction of a β-hydroxy ketone, where the β-hydroxy ketone comes from an aldol reaction. In the carbonyl condensation reaction, the carbonyl compounds behave both as nucleophile and as electrophile. In the presence of base, one carbonyl with α-hydrogen is converted into its enolate ion, which acts as nucleophile to attack the second carbonyl compound (the electrophile). The general condensation reaction for aldehydes and ketones with α-hydrogen is called aldol reaction. The aldol reaction is a reversible reaction (mechanism depicted in Figure 3.2). The equilibrium of reaction depends on both reaction conditions and the structure of substrates. A condensation product is generally favored with no α substituent (RCH₂CHO). A mixture of products will be generated from an aldol reaction of two similar aldehydes or ketones. However, if one of the carbonyl compounds contains no α-hydrogen but contains a non-steric hindered carbonyl group, such as benzaldehyde or formaldehyde, which would play a role as good electrophilic acceptor.
Figure 3.2. General mechanism of a typical aldol reaction
3.2 Materials and Methods

Organic solvents were purchased from Thermo Fisher Scientific (Waltham, MA, USA) otherwise noticed. All the chemicals were purchased from VWR (Radnor, PA, USA) otherwise noticed. Silica gel column chromatography was used for purification. Synthesized compounds were characterized by thin layer chromatography (TLC), $^1$H nuclear magnetic resonance (NMR), gas-chromatography-mass spectrometry (GC-MS) and high-resolution accurate mass (HRAM) mass spectrometry (MS). GC/MS spectra were acquired on Shimadzu QP5000 with DB5-MS capillary column with helium (He) as carrier gas. For monomeric compounds, following GC thermal gradient was used: initial temperature, 100°C, held for 5 min, increased from 100°C to 280°C at rate of 12°C/min, held for 10 min at 280°C. GC samples were derivatized with BSTFA prior to analysis by GC/MS. NMR spectra were acquired on a Varian Model NOVA400 (400 MHz) spectrometer. Acetone-$d_6$ was used as NMR solvent. High-resolution mass spectra were acquired by Q Exactive Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA) with electrospray ionization (ESI) in positive ion mode and negative ion mode. Positive ion mode: spray voltage: 3.8 kV, capillary temperature: 225°C, sheath gas: 2.0, probe heater temperature: 45°C. Negative ion mode: spray voltage: 3.2 kV, capillary temperature: 320°C, sheath gas: 10.0, probe heater temperature: 30°C, infusion flow rate for both modes: 5 µl/min. The syntheses of all the building block are shown in Figure 3.3

Synthesis of Compound 2

To a solution of vanillin (5 g, 32.9 mmol) in 50 ml acetone, 1.5 equiv. of ethyl bromoacetate (1M, 5.47 ml) and 1.5 equiv. of potassium carbonate (K$_2$CO$_3$) (6.8 g) were
added at room temperature. The reaction mixture was refluxed for 3 hours. TLC was used for monitoring the reaction progress. The reaction mixture was cooled to room temperature, and filtered. The filtrate was evaporated under vacuum to give clear oil residue. The residue was crystalized from ethanol to give pure 2 (6.88 g, yield 88%). $^1$H NMR (400 MHz, Acetone-$d_6$) $\delta$ 9.87 (s, 1H), 7.59 – 7.36 (m, 2H), 7.08 (d, $J = 8.2$ Hz, 1H), 4.87 (s, 2H), 4.22 (q, $J = 7.1$ Hz, 2H), 3.92 (s, 3H), 1.25 (t, $J = 7.1$ Hz, 3H).

**Synthesis of Compound 3**

To a solution of 2 (5 g, 21 mmol) in 125 ml toluene, 10 equiv. of ethylene glycol (11.7 ml) and catalytic amount of $p$-toluenesulfonic acid (TsOH) were added at room temperature. The reaction mixture was refluxed for 4 hours using a Dean-Stark apparatus to remove water. Thin layer chromatography (TLC) was used to monitor the reaction progress. The reaction mixture was cooled to room temperature, quenched with 5ml $K_2CO_3$ solution, extracted with ethyl acetate (EtOAc) (100 ml x 3), washed with brine and the organic layer was dried over $Na_2SO_4$, concentrated in vacuo to give light yellow residue. The residue was crystalized from ethanol to give pure 3 (5.5 g, yield 93%). $^1$H NMR (400 MHz, Acetone-$d_6$) $\delta$ 7.08 (d, $J = 1.9$ Hz, 1H), 6.97 (ddd, $J = 8.2$, 1.9, 0.5 Hz, 1H), 6.90 (d, $J = 8.2$ Hz, 1H), 5.67 (s, 1H), 4.71 (s, 2H), 4.20 (q, $J = 7.1$ Hz, 2H), 4.09 – 4.05 (m, 2H), 3.97 – 3.94 (m, 2H), 3.84 (s, 3H), 1.24 (t, $J = 7.1$ Hz, 3H).

**Synthesis of Compound 4**

To a solution of vanillin 1 (5 g, 32.9 mmol) and imidazole (4.4 equiv., 9.8 g) in 100 ml $CH_2Cl_2$, TipsCl (1.5 equiv. 10.5 ml) was added slowly. The reaction was stirred
Figure 3.3 The synthesis of the three building blocks for the synthesis of target trimer.
overnight, then extracted with CH₂Cl₂ (100 ml x 3), dried in vacuo to give colorless oil residue to white amorphous solid. The residue was purified with silica gel column to remove unreacted vanillin, mobile phase: EtOAc: Hexane= 1:5.¹H NMR (400 MHz, Acetone-\(d_6\)) \(\delta\) 9.88 (s, 1H), 7.53 – 7.41 (m, 2H), 7.08 (d, \(J = 8.5\) Hz, 1H), 3.92 (s, 3H), 1.31 (dq, \(J = 14.2, 7.4\) Hz, 3H), 1.11 (d, \(J = 7.4\) Hz, 18H). High-Resolution MS: [M+H]\(^+\) = 309.1877, Elemental composition: C\(_{17}\)H\(_{29}\)O\(_3\)Si, calculated: C\(_{17}\)H\(_{29}\)O\(_3\)Si, m/z 309.1886, \(\Delta\)ppm = -0.9820 ppm.

**Synthesis of Compound 5**

Compound 5 was synthesized according to the reference\(^{38}\). Briefly, 5 g of ferulic acid was dissolved in 50 ml ethanol, 2.5 ml acetyl bromide was added dropwisely. The reaction mixture was stirred overnight, extracted with 50 ml potassium carbonate and ethyl acetate (50 ml x 3). The organic layer was dried over Na\(_2\)SO\(_4\) and concentrated in vacuo to give light yellow oil. The solid was recrystallized in -20°C to give light yellow solid (5.2 g, yield 91%).¹H NMR (400 MHz, Acetone-\(d_6\)) \(\delta\) 8.12 (s, 1H), 7.59 (d, \(J = 15.9\) Hz, 1H), 7.33 (d, \(J = 2.0\) Hz, 1H), 7.14 (ddd, \(J = 8.1, 2.0, 0.5\) Hz, 1H), 6.87 (d, \(J = 8.2\) Hz, 1H), 6.39 (d, \(J = 15.9\) Hz, 1H), 4.19 (q, \(J = 7.1\) Hz, 2H), 3.92 (s, 3H), 1.27 (t, \(J = 7.1\) Hz, 3H).

**Synthesis of Compound 6**

To a solution of 5 (10 g, 51.5 mmol) in 200 ml acetone, 1.5 equiv. of ethyl bromoacetate (1M in THF, 8.57 ml) and 1.5 equiv. of potassium carbonate (K\(_2\)CO\(_3\)) (10.6 g) were added at room temperature. The reaction mixture was refluxed for 4 hours,
monitored using TLC, cooled to room temperature, and extracted with CH$_2$Cl$_2$ (100 ml x 3). The organic layer was dried with Na$_2$SO$_4$ and concentrated in vacuo to give light orange solid. The solid is recrystallized using THF in -20°C to give white solids (6.3 g, yield 90.7%). $^1$H NMR (400 MHz, Acetone-$d_6$) NMR (400 MHz, $J$ = 15.9 Hz, 1H), 7.35 (d, $J$ = 2.1 Hz, 1H), 7.20 – 7.12 (m, 1H), 6.93 (d, $J$ = 8.3 Hz, 1H), 6.44 (d, $J$ = 15.9 Hz, 1H), 4.77 (s, 2H), 4.20 (qd, $J$ = 7.1, 5.3 Hz, 4H), 3.91 (s, 3H), 1.30 – 1.26 (m, 3H), 1.26 – 1.22 (m, 3H). High-Resolution MS: [M+H]$^+$ = 309.1333, Elemental composition: C$_{16}$H$_{21}$O$_6$, calculated: C$_{16}$H$_{21}$O$_6$, m/z 309.1338, $\Delta$ppm= 0.2111 ppm.

**Synthesis of Compound 7**

To a solution of sodium bis(trimethylsilyl)amide (1.3 equiv. 6.9 g) in 55 ml anhydrous THF at -78°C under N$_2$, a solution of 3 in 18 ml anhydrous THF (7.73 g, 1.0 equiv) was added drop-wisely over 1 hour at -78°C. The reaction mixture was kept stirring for another 1hr at -78°C, then changed to -45°C. A solution of 4 in 20 ml anhydrous THF (10.13 g, 1.2 equiv) was added over 1hr at -45°C. After stirring for another 1hr, the reaction was quenched and extracted by addition of EtOAc (100 ml x 3) and K$_2$CO$_3$ solution. The organic layer was filtered through Na$_2$SO$_4$ and concentrated in vacuo to give light yellow oil residue (23 g, yield 142%). The crude oil residue was used without further purification. HRMS: [M+NH$_4$]$^+$ = 608.3256, Elemental composition: C$_{31}$H$_{50}$O$_9$NSi, calculated: C$_{31}$H$_{50}$O$_9$NSi, m/z 608.3256, $\Delta$ppm=1.027 ppm

**Synthesis of Compound 8**

To a solution of 7 (crude residue, 23 g, 4.0 equiv.) in 150 ml Acetone: H$_2$O (4:1, v: v), pyridinium p-toluenesulfonate (PPTS) (4.9 g, 1.0 equiv.) was added. The reaction
mixture was refluxed for 4 hours and cooled to room temperature. Acetone was removed in vacuo. The residue was extracted with EtOAc (100 ml x 3), washed with K$_2$CO$_3$. The organic layer was dried with Na$_2$SO$_4$ and concentrated under vacuum. The crude oil residue was used without further purification (7.21 g, crude yield 64%). HRMS: [M+NH$_4$]$^+$ = 564.2992, Elemental composition: C$_{29}$H$_{46}$O$_8$NSi, calculated: C$_{29}$H$_{46}$O$_8$NSi, m/z 564.2993, ∆ppm=0.832 ppm.

**Synthesis of Compound 9**

To a solution of 8 (5.7 g, 1.0 equiv) in 80 ml CH$_2$Cl$_2$, add imidazole (3.1 g, 4.4 equiv) and t-butyldimethylsilyl chloride (TBDMSCl) (3 g, 2.0 equiv). The reaction mixture was stirred overnight at room temperature. The reaction was quenched with 20 ml H$_2$O and extracted with EtOAc (50 ml x 3), washed with brine. The organic layer was concentrated in vacuo to give oil residue, purified by flash chromatography using silica gel, mobile phase: EtOAc/Hexane=1:8 give pure 9 consisted of the two diastereomers (3 g, 44%). White crystal formed in 4˚C. $^1$H NMR (400 MHz, Acetone-$d_6$) δ 9.83 (s, 1H), 7.43 – 7.38 (m, 2H), 7.20 (d, J = 2.0 Hz, 1H), 6.97 (s, 1H), 6.92 (d, J = 8.7 Hz, 1H), 6.86 (d, J = 8.1 Hz, 1H), 5.14 (d, J = 6.7 Hz, 1H), 4.81 (d, J = 6.7 Hz, 1H), 4.21 – 4.16 (m, 2H), 3.87 (s, 3H), 3.85 (s, 3H), 1.24 (t, J = 7.1 Hz, 6H), 1.08 (d, J = 7.3 Hz, 18H), 0.88 (s, 9H), 0.10 (s, 3H), -0.09 (s, 3H). HRMS: [M+NH$_4$]$^+$ = 678.3861, Elemental composition: C$_{35}$H$_{60}$O$_8$NSi$_2$, calculated: C$_{35}$H$_{60}$O$_8$NSi$_2$, m/z 678.3857, ∆ppm=1.346 ppm.

**Synthesis of Compound 10**

Solution of 6 (932 mg, 2.0 equiv in 12 ml anhydrous THF) was added dropwisely to a solution of NaHMDS (751 mg, 2.6 equiv in 60ml anhydrous THF) at -78˚C under N$_2$. 

89
After 30 min at -78°C, a solution of 9 (1 g, 1.0 equiv in 20 ml anhydrous THF) was added slowly. The reaction was quenched with saturated NH₄Cl after 15 min, extracted with EtOAc (100 ml x 3), washed with brine. The organic layer was dried over Na₂SO₄ and concentrated under vacuum to give yellow color residue. Compound 10 was purified by flash chromatography using silica gel, gradient mobile phase: EtOAc/Hexane 1:9 to 1:3. The purified 10 which consisted of a mixture of diastereomers were further separated (440 mg, yield 30%). HRMS: [M+NH₄]⁺ = 986.5114, Elemental composition: C₅₁H₈₀O₁₄NSi₂, calculated: C₅₁H₈₀O₁₄NSi₂, m/z 986.5117, Δppm=0.481 ppm.

**Synthesis of Compound 12**

To a solution of 10 (700 mg, 1.0 equiv) in 45 ml CH₂Cl₂ at -78°C under nitrogen, a solution of diisobutylaluminum hydride (DIBAl-H) (13.7 ml, 1M in THF, 19.0 equiv) was added dropwisely over 20 min. After 1.5 hours, the reaction mixture was brought to 0°C in ice-water bath, quenched by addition of 8 ml MeOH, extracted with EtOAc (100 ml x 3), washed with brine and dried through Na₂SO₄. The organic layer was concentrated to give viscous oil (596 mg, crude oil yield 98%). The crude residue was used without further purification. HRMS: [M+NH₄]⁺ = 860.4797, Elemental composition: C₄₅H₇₄O₁₁NSi₂, calculated: C₄₅H₇₄O₁₁NSi₂, m/z 860.4800, Δppm=0.219 ppm.

**Synthesis of Compound 13**

To a solution of 12 (crude oil, about 500 mg, 1.0 equiv) in 80 ml THF, 5 equiv of TBAF (3 ml, 1M in THF) was added at 0°C for 5 min, then switched to room temperature. The reaction is monitor with TLC. After 4 hours, THF was evaporated under vacuum.
Purification by flash chromatography, gradient mobile phase was used. 1) EtOAc with 5% TEA; 2) Acetone: EtOAc =1:9; and 3) Acetone: EtOAc= 1:4. (96.3 mg, yield 28.4%).

$^1$H NMR (400 MHz, Acetone-d6) $\delta$ 7.12 – 7.01 (m, 3H), 6.90 – 6.83 (m, 5H), 6.77 – 6.72 (m, 1H), 6.56 – 6.39 (m, 1H), 6.24 (dt, J = 15.8, 5.6 Hz, 1H), 4.94 – 4.79 (m, 2H), 4.38 – 4.21 (m, 2H), 4.16 (d, J = 3.6 Hz, 1H), 3.92 – 3.75 (m, 11H), 3.74 – 3.64 (m, 2H).

HRMS: $[M+\text{NH}_4]^+$ = 590.2598, Elemental composition: C$_{30}$H$_{40}$O$_{11}$N, calculated: C$_{30}$H$_{40}$O$_{11}$N, m/z 590.2601, $\Delta$ppm=0.411 ppm.

3.3 Results and Discussion

3.3.1 Synthesis of Building Blocks 3, 4 and 6

Compound 3 was obtained in high overall yield (81%) from vanillin in two straightforward steps: nucleophilic substitution with ethyl bromoacetate and continually protection of aldehyde group with ethylene glycol.

Compound 4 is to protect the free phenolic alcohol of vanillin. Traditional benzyl group is not used in this step because the terminating group 6 contains unsaturated side chain, which would also be reduced during the hydrogenolysis of the benzyl group with Pd/C and H$_2$. Silyl ether was found to be able to overcome this problem with reasonable yield and also it cut additional steps out in removing the protection group in later steps. Triisopropylsilyl (TIPS) was chosen to protect phenoxy group because it shows high stability in survival of aldol reaction compared with other silyl ether.

Compound 6 was obtained in overall satisfied yield from ferulic acid in two consecutive steps: first, esterification of ferulic acid with acetyl bromide to obtain
compound 5, the detailed information can be found in the previous report\textsuperscript{38}. Following step was the substitution of compound 5 with ethyl bromoacetate

\textbf{3.3.2 Synthesis of The Dimeric Intermediate}

As discussed previously, a dimeric intermediate containing $\beta$-hydroxy ketone was synthesized through aldol reaction. In the presence of strong base, the acidic $\alpha$-hydrogen in compound 3 was removed to generate the enolate ion as a nucleophile. Non-nucleophilic base was used so that the base will not compete with enolate ions in the aldol reaction as a nucleophile. Bases such as lithium diisopropylamide (LDA) or sodium bis(trimethylsilyl)amide (NaHMDS) with bulky substituents are poor nucleophile. Addition of the sodium enolate of compound 3 to compound 4 at -78°C in dry THF results in a mixture of diastereomers 7. Although TIPS group showed much better stability than trimethylsilyl (TMS) group, small amount of vanillin was observed resulting from loss of TIPS- from compound 4 in the presence of NaHMDS. This is due to the presence of H$_2$O with strong base in the reaction; even amount of water was only at trace level. To avoid the deprotection of compound 4, excess amount of NaHMDS was used to consume the residual water content. By this way, generation of vanillin during this step was successfully minimized. As no loss of TIPS from dimeric product 7 was observed, it indicated that TIPS on compound 7 was more stable than that on compound 4. The formation of $\beta$-hydroxyl group in compound 7 disrupted the conjugation system in TIPS-vanillin (compound 4) with aldehyde group contributing extra conjugation to the aromatic ring. As a result, vanillin may have more acid phenoxy group than compound 7, which leads to a less stable TIPS.
The mixture of diastereomers 7 was continually treated with PPTS to remove the 1,3-dioxolane protecting group, which led to compound 8. Bis(tert-butyldimethylsilyl) (TBDMS) ether was used directly to protect α-hydroxyl group. TBDMS provided a satisfied solution to protect β-hydroxyl group compared with acetal \(^{44}\) by reducing the number of steps that would be needed for protection and deprotection. The protected dimeric intermediate 9 formed nice white crystal in the mixtures of hexane and ethyl acetate. The X-ray crystal structure of compound 9 was collected by Dr. Sean Parkin from department of Chemistry at University of Kentucky (shown in Figure 3.4).

### 3.3.3 Synthesis of The Trimeric Model

Similar with previous dimer synthesis, trimeric compound 13 was synthesized via aldol reaction using dimeric intermediate unit 9 and terminating unit 6. Compound 6 was first treated with strong base to generate enolate ion. LDA is a very common non-nucleophile base used in aldol reaction. The terminating unit 6 contains a \(\alpha\), \(\beta\)-unsaturated ester group. The electronegative oxygen from carbonyl group causes \(\beta\) carbon to be more electrophilic than a typical alkene carbon and results in a resonance stabilized enolate ion. Because of additional reactivity of 6, the resulting enolate ion can react with different components leading to un-wanted by-product. The possible reaction pathways are shown in Figure 3.5.a.
Figure 3.4 X-ray crystal structure of compound 9
Figure 3.5. a) Reaction pathway of nucleophilic addition of di-isopropanol amine to compound 6 (purple color); nucleophilic addition within compound 6 (blue color); b) Possible byproduct 10 resulting from elimination.
Due to the steric effect, conjugate addition of enolate ion to β carbon on side chain (Figure 3.5.a, pathway showed in blue trace) is less favored than addition to the aldehyde group on dimer 9 (Figure 3.5.a, pathway showed in red trace). The trimeric product is preferred than a self-conjugation addition of compound 6 (Figure 3.5a, blue color). However, the predominant by-products with addition of diisopropylamide group, compound 14 was observed\(^{47}\) in the coupling reaction. Although LDA is considered to be a non-nucleophilic base, the conjugate addition of diisopropyl amide to α, β-unsaturated carbonyl group of compound 6 occurred resulting in forming predominate by-product of compound 14\(^{47}\) (Figure 3.5a pathway in pink color). Thus, alternative organic bases that contain bulkier group are considered for this reaction. NaHMDS turned out to be a satisfied non-nucleophilic base to overcome this problem. Due to the steric effect of the bulky bis(trimethylsilyl)amide (HMDS) group, there was no conjugate addition of HMDS to β-carbon observed.

No loss of TIPS from compound 9 was observed in the later nucleophilic addition with compound 6. As discussed previously, that the disrupting conjugation system of compound 9 led to a more stable TIPS protection. The residue from this step required no purification for the use in the following step.

Finer separation of diastereomers can be achieved by using silica gel column. The reduction of the ethyl ester group of compound 10 with diisobutylaluminum hydride (DIBAL) at -78°C afforded good yield. The reaction went to completion within an hour. The reaction temperature and work-up procedure were crucial in obtaining quantitative yield. In this step, undesired by-product 10 formed in higher reaction temperature, which was resulted from the elimination of TBDMS group of dimeric compound 9 in the basic
condition (structure shown in Figure 3.5.b). By keeping reaction temperature at -78°C, one can quantitatively reduce compound 10 in an hour. During work-up process, it was found that the reaction needed to be warmed up near 0°C before quenching with methanol, otherwise, the target products could convert to other by-products (structures unknown) immediately. The mechanism of this conversion was unclear, but these byproducts have higher molecular weight that target trimer. Since silyl ether (TIPS and TBDMS) was used as protecting group for both phenolic and β-hydroxyl groups, they can be removed under the same condition at the same time with tetrabutylammonium fluoride (TBAF) in THF. It was important to keep this order of reduction and deprotection in order to obtain better overall yield and simplify the additional purification process. Each procedure involving the synthesis of trimer was pretty clean reaction, and there is only one purification step needed for the target trimeric β-O-4 compound 13. Figure 3.6 shows the mass spectrum of deprotonated compound 13 in the negative ion mode as well as its HCD/MS/MS spectrum.
Figure 3.6. a) ESI-MS of targeted trimer in negative ion mode, [M-H]^−; b) HCD/MS/MS spectrum of the trimer, [M-H]^−.
3.4 Conclusion

A trimeric β–O-4 lignin model compound is synthesized in 7 steps from vanillin and ferulic acid, which gives a repeating guaiacyl unit in the trimeric compound. The purpose of this synthesis route is to incorporate additional α, β-unsaturated side chain to β-O-4 lignin. This new trimeric compound includes phenoxy end group, β-O-4 moiety and α, β-unsaturated side chain, which will be an interesting model to study its physical and chemical properties. No previous work has been done in introducing different functionalities including unsaturated side chain in the synthesis of lignin model compound using this strategy. Considering the reactivity of α, β-unsaturated side chain in the terminating group, new organic base and protecting group have been investigated to replace LDA and benzyl bromide, which are commonly used in synthesizing β-O-4 linkage model compounds. In the procedure developed in this chapter, silyl ether as protecting group was used for phenolic and aliphatic alcohol. Silyl ether showed very good stability under conditions containing NaHMDS and DIBAl-H. As both alcohols were protected with silyl ether, all these silyl ethers for different alcohol groups could be removed at the same time, which reduces of the reaction steps. This synthesis procedure is also feasible for preparing higher order oligomers and incorporating different monomeric units, such as syringyl (S) and 4-hydroxyphenyl (H).
4.1 Introduction

4.1.1 Pretreatment Strategies

Lignocellulosic biomass is quite complex, composed of cellulose, hemicellulose and lignin. As a result, plants are quite resilient to degradation under normal conditions. Naturally, biomass can undergo decomposition by microbes in the moist conditions, however natural degradations are usually quite slow and may take several years to complete. In order to use biomass in biofuel production, a wide variety of pretreatment techniques have been studied to break the hemicellulose-lignin complex cross-links and increase the rate of biomass-biofuel conversion, shown in Figure 4.1.

There are generally two categories of techniques for conversion of biomass to biofuels: thermochemical conversion and biochemical conversion. Biochemical conversion techniques require much lower temperatures compared with thermochemical conversion. As discussed previously that the natural properties of lignin, such as, aromaticity, extensive carbon-carbon cross-linkage, and structural heterogeneity make lignin the most recalcitrant in protecting biomass from decomposition. As a result, the presence of lignin largely hinders the accessibility of cellulose and other polysaccharides to enzymatic hydrolysis step during the biochemical conversion. The first key step in biochemical conversion is the pretreatment of lignocellulosic biomass. The goal of
Figure 4.1. Schematic of pretreatment step of biomass
pretreatment is to change the physical and chemical properties of biomass material (surface area, cellulose crystallinity index, degree of polymerization, lignin content and etc.) to increase the rate of enzymes hydrolysis.

Pretreatment methods can be divided into several groups in general: mechanical (milling and grinding), physicochemical (steam pretreatment, hydrothermolysis and wet oxidation), chemical (alkali, dilute acid, oxidation and organic solvents), and biological (fungi)\textsuperscript{48}.

The purpose of physical pretreatments is to reduce the particle size and increase the surface area of biomass. For instance, disk milling/ball milling/grinding usually decrease the size of biomass to 0.2-2 mm size. But the milling process normally requires intensive energy input. Chemical pretreatments offer a wide variety of pretreatment methods and conditions. The pretreatment can be carried out at acidic, neutral or basic conditions. For acidic conditions, such as mineral acids (H\textsubscript{2}SO\textsubscript{4}, HCl and HNO\textsubscript{3}) or organic acids (formic acid, acetic acid and maleic acid), the pretreatment temperature, pressure and concentration of acids largely affect the product compositions. Alkaline pretreatments usually swell and increase the surface area of cellulose, disrupt the crystallinity of cellulose and hydrolyze the lignin-carbohydrate cross-linkage, which leave hemicellulose and cellulose behind. Strong alkali species, such as NaOH and KOH can cleave ester and ether bonds, while weak base like ammonia cleaves only ester bonds. Table 4.1 is the brief summary of the conditions, advantages and disadvantage of different pretreatment methods\textsuperscript{4, 48}. 

\begin{center}
\begin{tabular}{|c|c|c|}
\hline
 Method & Conditions & Advantages & Disadvantages \\
\hline
Mechanical & High particle size & High energy input & \textless \textbar \textbf{Mechanical} \\
\hline
Physicochemical & Steam & Efficient & \textless \textbf{Physicochemical} \\
\hline
Chemical & Alkaline & Efficient & \textless \textbf{Chemical} \\
\hline
Biological & Fungi & Efficient & \textless \textbf{Biological} \\
\hline
\end{tabular}
\end{center}
Table 4.1. Summary of various pretreatment strategies

<table>
<thead>
<tr>
<th>Pretreatment</th>
<th>Chemicals Needed</th>
<th>Post pretreatment Solid</th>
<th>Post pretreatment liquid</th>
<th>Pretreatment conditions</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical</td>
<td>Disk milling</td>
<td>NA</td>
<td>Whole Biomass</td>
<td>NA</td>
<td>Milling (10-30mm) Grinding, particle size (0.2-2µm)</td>
<td>No chemicals required, scalable, no water used</td>
</tr>
<tr>
<td>Acidic</td>
<td>Dilute Sulfuric acid</td>
<td>Dilute sulfuric acid</td>
<td>Enriched cellulose</td>
<td>Xylose</td>
<td>140-190 °C, 0.4-2% sulfuric acid, 1-40 min</td>
<td>Applicable to wide variety of materials, Produce hydrolyzed xylose during pretreatment</td>
</tr>
<tr>
<td>Organic acid</td>
<td>Acetic acid, fumaric acid, formic acid</td>
<td>Acetic acid, fumaric acid, formic acid</td>
<td>Enriched cellulose</td>
<td>Soluble hemicellulose and lignin</td>
<td>130-190 °C, 50-90 mM of organic acid</td>
<td>Separation of biomass to lignin and hemicellulose rich fraction and cellulose rich fraction, low pressure</td>
</tr>
<tr>
<td>Concentrated acid</td>
<td>H₂SO₄, HF, HCl, H₃PO₄, HNO₃</td>
<td>Condensed lignin</td>
<td>Soluble glucose</td>
<td>Short reaction time</td>
<td>Hydrolysis of cellulose during pretreatment, largely decrease the crystallinity of cellulose to amorphous cellulose, effective on soft wood</td>
<td>Hydrolysis of cellulose during pretreatment, largely decrease the crystallinity of cellulose to amorphous cellulose, effective on soft wood</td>
</tr>
<tr>
<td>Acidic organosolv</td>
<td>Methanol, ethanol, acetone, ethylene glycol and tetrahydrofurfuryl alcohol, water</td>
<td>Enriched cellulose and majority of the hemicellulose</td>
<td>Lignin and some soluble hemicellulose</td>
<td>Acetone-water (1:1 ratio) pretreatment at 195°C, pH 2.0, high pressure required</td>
<td>Fractionation of lignin stream, cellulose digestibility increases due to</td>
<td>High risk of dealing with high-pressure reactor, flammable and volatile organic solvents</td>
</tr>
<tr>
<td>Pretreatment Type</td>
<td>Ionic Liquid</td>
<td>Neutral and Alkaline</td>
<td>Biological</td>
<td>Physiochemical</td>
<td>Steam Explosion</td>
<td>Fermentation with Fungi or Bacteria</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------</td>
<td>----------------------</td>
<td>------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td><strong>Ionic Liquid</strong></td>
<td>1-ally-3-methylimidazolium-chloride ([AMIM]Cl), 1-ethyl-3-methylimidazolium-acetate ([EMIM]Ac).</td>
<td>Enriched Cellulose and hemicellulose content</td>
<td>Lignin and some hemicellulose</td>
<td>100-150°C, several min to hours</td>
<td>Low in the loss of carbohydrate, few degradation products except at severe conditions</td>
<td>High solvent cost, high cost in regeneration of ionic liquid.</td>
</tr>
<tr>
<td><strong>Ammonium fiber expansion (AFEX)</strong></td>
<td>Liquid or gaseous anhydrous ammonia</td>
<td>Whole biomass</td>
<td>NA</td>
<td>100-140°C, 1:1- 2:1 ammonia to biomass loading, 30-60 min residence time, 60-100% moisture content</td>
<td>Ammonia can be recovered and reusable, few degradation products, dry process</td>
<td>Safety precautions for dealing with volatile ammonia, not efficient for hardwood biomass</td>
</tr>
<tr>
<td><strong>NaOH</strong></td>
<td>NaOH</td>
<td>Amorphous cellulose and some hemicellulose</td>
<td>Soluble lignin and hemicellulose</td>
<td>NaOH</td>
<td>Decrease the crystallinity of cellulose to highly reactive amorphous cellulose, Solubilized lignin</td>
<td>Longer residence time required, large amount of water used, difficult in scaling up, expensive recovery of NaOH</td>
</tr>
<tr>
<td><strong>Lime</strong></td>
<td>Calcium monoxide with and without oxygen</td>
<td>Whole biomass</td>
<td>NA</td>
<td>25-160°C, 120min to weeks, 0.07-0.2 g CaO/g biomass</td>
<td>Inexpensive pretreatment reactor system</td>
<td>Large amount of water used, high cost in recovery of catalyst, long residence time</td>
</tr>
<tr>
<td><strong>Alkaline peroxide</strong></td>
<td>NaOH, H₂O₂</td>
<td>Enriched cellulose and some hemicellulose</td>
<td>Soluble Degraded lignin and hemicellulose</td>
<td>0.5-2% NaOH, 0.125 g H₂O₂/g biomass, 22°C, atmospheric pressure for 2 days</td>
<td>Mild pretreatment condition, commercially used in paper industry</td>
<td>Large amount of water used, high cost in catalyst recovery, loss of lignin due to oxidation</td>
</tr>
<tr>
<td><strong>Physiochemical</strong></td>
<td>Steam and SO₂</td>
<td>Enriched cellulose</td>
<td>Soluble hemicellulose</td>
<td>180-210°C, 1-120min</td>
<td>Suitable for hardwood and herbaceous biomass</td>
<td>High cost in reactor due to high pressure operation</td>
</tr>
<tr>
<td><strong>Biological</strong></td>
<td>Microbes like fungi and bacteria</td>
<td>Whole biomass with loss of cellulose and hemicellulose</td>
<td>NA</td>
<td>25-30°C solid state fermentation, 80-120% moisture content, 10-15 days</td>
<td>Mild pretreatment condition, low energy required</td>
<td>Slow process, not very high in sugar conversion, need continuous monitoring</td>
</tr>
</tbody>
</table>
Biological pretreatment methods usually require less expensive reaction systems compared with chemical or physical pretreatment process. Basidiomycetes has drawn a lot of interest in application of biodegradation of lignocellulose, as it plays an important role in decomposing woody material for carbon cycling in the ecosystem\(^{49}\). There are two different groups of lignocellulolytic Basidiomycetes: white-rot fungi and brown-rot fungi. Brown-rot fungi usually decompose wood by degrading/removing carbohydrates and leaving lignin as modified without further degradation. White-rot fungi, such as *Phaneochaete chrysosporium*, show ability to mineralize lignin, open up cell walls and increase carbohydrate accessibility\(^{50}\). The studies have shown that highly oxidative hydroxyl radical (OH), generated by extracellular Fenton chemistry (Fe\(^{2+}\) and H\(_2\)O\(_2\)), causes the destruction of lignin at close proximity\(^{49}\). Lignin is likely to undergo depolymerization and re-polymerization forming different chemical structures, which overall results in the increasing accessibility to the polysaccharides in the cell wall.

However, biological pretreatment is usually a time-intensive process with lower throughput. In order to mimic white-rot fungi degradation of lignocellulosic biomass, solution phase Fenton chemistry opens up possibilities to generate hydroxyl radical and non-selectively oxidize organic species in the *in-vitro* condition. Typically, iron serves as an electron donor to hydrogen peroxide, which causes the catalytic decomposition of hydrogen peroxide in the generation of hydroxyl radicals. The classic Fenton’s reaction has been widely used in industrial wastewater treatment for degradation of toxic organics, decomposition of pesticides and herbicides\(^{51-52}\). Only a few studies apply the chemical Fenton’s reaction onto the pretreatment of plant material. Because the simplicity, low cost of Fenton’s reaction, it provides potential application for high
throughput and efficient pretreatment in biomass-to-biofuel conversion compared with biochemical Fenton chemistry using white fungi. Fenton’s reaction has shown the ability to open up plant cell wall and increase cellulose accessibility on various biomass feedstock\textsuperscript{53}. However, previous analysis on the Fenton-pretreated biomass shows that there is no lignin content difference in the post-treated biomass. No lignin-like degradation products were detected in the supernatant of pretreatment solution by Fenton’s reaction as well.

4.1.2 Untargeted analysis by High-resolution Accurate Mass (HRAM) Mass Spectrometry

The types of degradation reaction involved in lignin breakdown highly depend on the reactive functional groups within the molecules. As discussed in the previous chapter, lignin contains several important types of chemical groups: 1) aromatic rings are chemically stable and ring opening is usually achieved by inserting oxygen onto ring to yield carboxylic acid by enzymes like intra-molecular dioxygenases\textsuperscript{54}; 2) ether linkages, such β-O-4, are major linkage in lignin and are stable and resistant to acid or base hydrolysis; 3) primary hydroxyl group on the γ-C and the secondary alcohol on the α-C resulted from dominant β-O-4 group; the free phenolic group is at a relative lower content because it extensively involves in lignin polymerization reaction. All these hydroxyl groups have potential to be oxidized to form carboxylic acid, ketone or quinone groups in the biochemical reactions\textsuperscript{54}.

\[
\text{Fe}^{2+} + \text{H}_2\text{O}_2 \rightarrow \text{Fe}^{3+} + \cdot \text{OH} + \text{OH}^- \quad (1) \quad \text{\textsuperscript{49}}
\]
The hydroxyl radical generated from Fenton chemistry is highly reactive and can randomly attack organic species at close proximity. As the random and non-specific properties of reactive hydroxyl radical, the innovative aspect of this project is to study the oxidation profile of the products from Fenton’s reaction in an untargeted strategy using bioinformatics tools to screen out the possible lignin related product features. High-performance liquid chromatography (HPLC) coupled with high-resolution mass spectrometry (HRMS) provides additional separation step for analyzing complex sample mixture, together with mass-to-charge and ion intensity information. Thus, handling such complex HPLC-HRMS datasets is a critical step for identification and quantification of compounds of interest from complex sample systems. Many bioinformatics tools have been developed in order to handle such complex datasets for proteomics and metabolomics studies, such as XCMS\textsuperscript{55} and MZmine\textsuperscript{2}\textsuperscript{56}. Those bioinformatics tools are widely used in data processing and analysis of metabolomics studies for detection of biologically meaningful metabolites as well as their dynamics related to environmental changes. Such bioinformatics tools are consist of a typical data processing platform, including several stages: filtering, detection of features, alignment and normalization\textsuperscript{57}. Filtering aims to remove interference signals, such as matrix backgrounds or random noise based on user define parameters. Feature detection step is to detect all the possible signals that are caused by true ions. In this step, both m/z and retention time information are pulled out to define chromatographic peak. During feature detection, isotopic pattern comparison can also be used to improve the detection accuracy by reducing the number of false positive assignments. Feature detection step is critical in accurately quantify the concentration of ions. It is common that retention time variation happens in LC
techniques. Alignment step is used for correcting the drifting in retention time between different runs and samples. The aim of normalization is to eliminate the undesired systematic deviation in ion intensities between runs and samples. This can be done by addition of internal standard compounds in the samples\(^{57}\).

In this study, high-resolution mass spectrometer, Thermo Scientific Q-Exactive Orbitrap, is used as mass detector coupled with HPLC for analysis of complex post-pretreatment samples. HRMS measurements provide several advantages in feature detection as well as identification: 1) high sensitivity for identifying low abundant species because of high mass accuracy and 2) assignation of molecular formula with high accurate mass data. HRMS is able to measure accurate mass of analyte within ppm accuracy range. Additionally, high-resolution mass measurement allows highly resolved isotopic peaks, which largely increases the confidence in determining the elemental composition unambiguously. Further analysis, including isotope pattern, tandem mass spectrometry enables the selection and identification of ions of interest.

In this chapter, synthetic trimeric model compound was used for the pretreatment study. Because of the complexity of native lignin, it is extremely difficult to understand Fenton chemistry at the molecular level when dealing with natural lignin. The problem is much less complicated when working with lignin model compounds. The trimer model used in this work was synthesized through a series of steps described in Chapter 3. As discussed, the trimeric model compound used in this chapter contains several important function groups, which will be a good representation for lignin as different functional groups may have different reactivity during the pretreatments. As discussed previously that the non-specificity and randomness in destructive reaction property of strong oxidant
-hydroxyl radicals in the Fenton chemistry, the product profile of post-oxidation with hydroxyl radical is expected to be challenging in complete prediction. Thus, the analytical strategy to investigate the product profile is to apply the similar untargeted analysis strategy in attempt to seeking and identifying the degradation products of trilignol model compound. The study will provide further understanding regarding the behavior of trimeric lignin molecules during the degradation process of Fenton chemistry and provide clues to help explain the observation during in the pretreatment of biomass feedstock with Fenton’s reaction.

4.2 Materials and methods

Materials

All chemicals and reagents were obtained and used without further purification. Hydrogen peroxide (H₂O₂, 50% wt.), ferrous chloride tetrahydrate (reagent grade), sodium thiosulfate, LC/MS grade of water, acetonitrile and methanol were purchased from Thermo Fisher Scientific (Waltham, MA, USA). Vanillin, ferulic acid, para-phenylphenol, coniferyl aldehyde, and 3-hydroxy-4-methoxybenzoic acid (Isovanillic acid) were purchased from Sigma Aldrich (Milwaukee, WI, USA). Coniferyl alcohol was synthesized in house from coniferyl aldehyde using published methods. Trilignolic model compound [G(β-O-4) G(β-O-4)G] (GGG) was synthesized as described in Chapter 3 that consisted with a mixture of diastereomers. A Kromasil® EternityXT Hexylphenyl column (2.1X100mm, 2.5μm particle size) was purchased from Sigma Aldrich (Milwaukee, WI, USA).
Fenton’s reaction pretreatment on trilignol

The GGG trimer was dissolved in 200 µl ethanol (EtOH) and then into 10ml distilled water (18 MΩ), followed by 250 µl, 50 mM FeCl₂ and 25.8 µl 50% H₂O₂ resulting in [GGG]₀ = 1 mM, [H₂O₂]₀ = 50 mM and [Fe²⁺]₀ = 1.25 mM. The reaction was stirred at room temperature. Reaction samples (200 µl) were taken at the following time points: 0 min, 5 min, 15 min, 35 min, 75 min, 155 min, and 315 min. The reaction samples were quenched with 0.2 mL, 0.1M Na₂S₂O₃ (aq), acidified with 0.2 ml, 1N HCl (aq), and extracted with 200 µL ethyl acetate containing 0.25 µg/ml 4-phenylphenol as internal standard. This reaction was done in three replicates. 100 µL of above ethyl acetate layer was taken, dried under N₂. The residue was then dissolved in 5% ACN, 5% THF and 90% H₂O for LC-MS analysis.

Sample Preparation for HPLC-HRMS analysis

For HPLC/MS analysis, a standard solution of 20µg/ml of mixed model compounds (vanillin, ferulic acid, isovanillic acid, coniferyl alcohol and synthesized trilignol GGG) was made in 95% H₂O, 5% ACN. The mixed standard solution was used to develop LC separation method for analysis of post Fenton pretreatment samples. A Kromasil® EternityXT HexylPhenyl column (100 x 2.1mm, 2.5µm) was used for all HPLC separations, at a flow rate of 0.3 ml/min. Mobile phase: A: H₂O, B: ACN. The eluent was introduced into Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific) by electrospray ionization (ESI) in the negative ion mode. Raw MS data were processed by MZmine2 software package. Briefly, after importing raw data to MZmine2, mass detection function was first used to detect individual ions in each scan.
and then a list of ions for each scan was created, followed by chromatogram builder step. In the chromatogram builder step, a chromatogram was reconstructed for each mass which was previously detected. Each chromatogram built previously was then deconvoluted into individual peaks. After grouping isotopic peaks and filtering the ions based on the given restriction. A final peak list contains all features was generated, including m/z, retention time and peak area. The peak area of each feature was manually integrated using Xcalibur and normalized to internal standard.

4.3 Results and Discussion

4.3.1 High-Performance Liquid Chromatography (HPLC) Method Development

To develop and evaluate the HPLC/MS methods for post Fenton pretreatment products analysis, mixture of pure standard compounds were used and separated by a reversed phase Hexylphenyl column. Hexylphenyl column was chosen because of its good efficiency in separating aromatic compounds. Several mobile phase gradients were tested with standard compounds. In order to be compatible with negative ion mode ESI, formic acid was not added to the mobile phase. The Kromasil® EternityXT Hexylphenyl column is packed with porous silica microspheres, which are functionalized with hexylphenyl stationary phase. After comparison of different gradients, the optimal nonlinear gradient elution is used for this column at a flow rate of 0.3ml/min (see Table 4.2).

Figure 4.2.b-e. show the EIC of each deprotonated model compounds. Vanillic acid, ferulic acid and coniferyl alcohol are well separated from each other. Peaks of
vanillin and coniferyl alcohol are very close to each other. With accurate mass EIC, these two peaks are able to be distinguished from each other. Figure 4.3 shows the EIC of the deprotonated trilignol (GGG). There are four peaks accounted for trilignol. The mixture of diastereomers are not able to be completely resolved in this LC condition. Since only the relative quantification of total amount of trilignol is interested in this study, it is acceptable that the diastereomers of trilignol are not fully separated from each other.
Table 4.2. LC method used in the pretreatment study

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>%A: H₂O</th>
<th>%B: ACN</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>5.00</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>20.00</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>23.00</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>25.00</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>27.00</td>
<td>90</td>
<td>10</td>
</tr>
</tbody>
</table>
Figure 4.2. Liquid chromatogram of mixture of standards; a) Base peak chromatogram (BPC); b) Extracted Ion Chromatogram (EIC) of 3-hydroxy-4-methoxybenzoic acid; c) EIC of ferulic acid; d) EIC of coniferyl alcohol; e) EIC of vanillin.
Figure 4.3. Extracted Ion Chromatogram (EIC) of starting trilignol GGG.
4.3.2 Analysis of Unknown Products from Fenton’s Reaction of Trilignolic Model Compound (GGG)

Thermo Scientific Q-Exactive Orbitrap mass spectrometer was coupled with HPLC system for high resolution MS data acquisition. All the standard model compounds formed stable deprotonated ions [M-H]. The high-resolution accurate mass data acquisition allows the determination of the elemental compositions for unknown post-pretreatment products as well as the standard compounds. To obtain structural information, data-dependent tandem mass spectrometric experiment (MS/MS) was performed. In data-dependent MS/MS experiment, one full scan at high-resolution (R=70,000) was carried out to determine the subsequent MS/MS experiment. Five most abundant ions in each full scan were then subjected to high-energy collision dissociation (HCD) cell for MS/MS experiment. The products ions were analyzed at a resolution of 17,500 in the Orbitrap mass analyzer. MS/MS experiments and high-resolution mass measurements can be used to propose the molecular structures to the detected unknown compounds. Upon MS/MS experiments, the deprotonated molecules of lignin related compounds share common fragmentation patterns, which indicate the existing of particular functional groups and chemical linkages\(^{30-31, \, 60}\). For example, loss of 15Da, 18Da, and 30Da are corresponding to the loss of methyl radical, water and formaldehyde group which indicates the presence of methoxy group on the aromatic ring and aliphatic alcohol group at \(\gamma\)-C, respectively. Elemental composition assignment with high-resolution analysis provides another level of confidence in interpreting the MS/MS spectrum as well.
Para-phenylphenol (PPP) was used as the internal standards for all the LC/MS analysis. PPP is chosen because of its similar aromaticity and phenoxy group, which gives reasonable response in the negative ion mode. PPP also showed good stability during the work-up process of solution phase Fenton chemistry. The condition of Fenton’s reaction used in this study was modified based a previous publish work\textsuperscript{52,61}. In this study, initial concentrations of three reagents were used as following: \([GGG]_0 = 1\) mM, \([H_2O_2]_0 = 50\) mM and \([Fe^{2+}]_0 = 1.25\) mM. Reaction mixtures were sampled at fixed intervals.

Data processing was performed using MZmine2 software\textsuperscript{58}. The peak list generated by MZmine2 contains all the features with m/z value, retention time and the peak area information. All the features were manually screened, based on the mass defect and MS/MS information. Table 4.3 is a summary of twenty features that have reasonable mass defect to be potential lignin-like compounds. The time course measurements of sixteen features were used to plot the concentration change curves. Figure 4.4 shows the concentration changes of starting trilignol over 5 hours.

As shown in Figure 4.4, there is a dramatically decreasing of GGG in the first 5 min, the degradation rate drops as the amount of the remaining GGG decreases. When there is only GGG and FeCl\textsubscript{2} presence in the reaction mixture at \(t = 0\) min, the generation of unknown peaks, which are around trimeric and tetrameric molecular weight ranges, are observed. The mechanism of production of trimeric and higher MW molecules is still unclear. Figure 4.5 shows the time course measurement of major trimeric unknown product \([M-H]^- 551.1909 (C_{30}H_{31}O_{10}, -1.6784\) ppm).
Table 4.3. List of 20 features with their corresponding elemental composition and retention time based on the result from MZmine 2.

<table>
<thead>
<tr>
<th>m/z</th>
<th>[M-H]^−</th>
<th>∆PPM</th>
<th>rt(min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>745.3051</td>
<td>C38H49O15</td>
<td>-1.4951</td>
<td>12.97</td>
</tr>
<tr>
<td>629.2594</td>
<td>C33H41O12</td>
<td>-1.3164</td>
<td>11.09</td>
</tr>
<tr>
<td>593.1428</td>
<td>C16H33O23</td>
<td>0.986</td>
<td>12.43</td>
</tr>
<tr>
<td>571.2179</td>
<td>C30H35O11</td>
<td>-1.0321</td>
<td>10.82</td>
</tr>
<tr>
<td>569.2023</td>
<td>C30H33O11</td>
<td>-0.8845</td>
<td>11.43</td>
</tr>
<tr>
<td>563.2131</td>
<td>C28H35O12</td>
<td>-1.3373</td>
<td>12.39</td>
</tr>
<tr>
<td>553.2064</td>
<td>C30H33O10</td>
<td>-0.6143</td>
<td>12.44</td>
</tr>
<tr>
<td>551.1909</td>
<td>C30H31O10</td>
<td>-1.6784</td>
<td>12.78</td>
</tr>
<tr>
<td>523.1597</td>
<td>C28H27O10</td>
<td>-2.6059</td>
<td>13.23</td>
</tr>
<tr>
<td>433.1856</td>
<td>C23H29O8</td>
<td>-3.0131</td>
<td>12.46</td>
</tr>
<tr>
<td>375.1439</td>
<td>C20H23O7</td>
<td>-2.68</td>
<td>12.17</td>
</tr>
<tr>
<td>375.1439</td>
<td>C20H23O7</td>
<td>-2.68</td>
<td>10.51</td>
</tr>
<tr>
<td>373.1282</td>
<td>C20H21O7</td>
<td>-2.8727</td>
<td>11.43</td>
</tr>
<tr>
<td>373.1282</td>
<td>C20H21O7</td>
<td>-2.3002</td>
<td>12.8</td>
</tr>
<tr>
<td>353.1597</td>
<td>C18H25O7</td>
<td>-2.745</td>
<td>12.98</td>
</tr>
<tr>
<td>213.0762</td>
<td>C10H13O5</td>
<td>-2.523</td>
<td>0.94</td>
</tr>
<tr>
<td>209.0812</td>
<td>C11H13O4</td>
<td>-2.9563</td>
<td>13.05</td>
</tr>
<tr>
<td>209.0812</td>
<td>C11H13O4</td>
<td>-3.0293</td>
<td>14.35</td>
</tr>
<tr>
<td>197.0449</td>
<td>C9H9O5</td>
<td>-3.4031</td>
<td>0.94</td>
</tr>
<tr>
<td>195.0292</td>
<td>C9H7O5</td>
<td>-3.6228</td>
<td>0.6</td>
</tr>
<tr>
<td>151.0397</td>
<td>C8H7O3</td>
<td>-2.5113</td>
<td>8.93</td>
</tr>
<tr>
<td>167.0346</td>
<td>C8H7O4</td>
<td>-2.0628</td>
<td>1.71</td>
</tr>
</tbody>
</table>
Figure 4.4. Time course measurement of starting material GGG. Error bars indicate ± one standard deviation.
Figure 4.5. Degradation of trimeric unknown feature, [M-H] 551.1909. Error bars indicate ± one standard deviation.
Similar to degradation pattern of trimer GGG, upon the addition of $\text{H}_2\text{O}_2$, trimeric compound ([M-H] $551.1909$) is subjected to degradation rapidly in the first 5 min and goes to completion less than 3 hours (Figure 4.5). Two possible structures of unknown $551.1909$ is proposed according the fragmentation pattern as well as the elemental composition of fragment ions (shown in Figure 4.6). The fragmentation pattern of ion $551.1909$ is very similar to original trimer GGG. They both contain fragment ion $195.0656$ with the same elemental composition, which indicates that the unknown trimer with m/z $551.1909$ contains same phenolic end group as in trimer GGG. Figure 4.7 depicts the possible fragmentation pathways of the putative structures that will lead to the fragmentation ions, which were observed in the HCD/MS/MS spectrum (Figure 4.6.a).

**Time Course Curve of Other Unknown Features**

The peak area of individual features at different time point was extracted out manually using Xcalibur. The time course measurements of a few unknown features are showed in Figure 4.8. The generation of the unknown features with monomeric and dimeric molecular weights are observed as Fenton reaction proceeds, such as [M-H]$^-$ $151.0397$, $375.1489$, and $373.1282$. Concentration of several dimeric and monomeric unknown compounds, such as m/z $375.1439$ (12.17 min) and m/z $151.0397$, increases as the reaction started and decreased after about 1hr, which indicates a further degradation occurs among these intermediate products as shown in Figure 4.8.
Figure 4.6. a) HCD/MS/MS spectrum of m/z 551.1909; b) putative structures of m/z 551.1909.
Figure 4.7. possible fragmentation structures for corresponding trimeric unknown m/z 551.1909
Figure 4.8. Time course plots of fifteen unknown features. Error bars indicate ± one standard deviation.
Identification of unknown features

Among these unknown features, three of them are identified and confirmed with standard compounds. Figure 4.9. shows the extracted ion chromatogram of [M-H]− with exact m/z 151.0397 at retention time of 8.96 min, corresponding to an elemental composition of C8H7O3, Δppm= -2.5 ppm and a double bond equivalent (DBE) of 5, which can be a compound with one aromatic ring and one double bond. Upon HCD/MS/MS experiment, this unknown m/z 151.0397 is fragmented via a methyl radical loss (15 Da) and gives rise to m/z 136.0161 (C7H4O3, Δppm= -3.6341), which indicates there is a methoxy group on the aromatic ring\textsuperscript{31}. The exact m/z value, elemental composition (EC) assignment as well as the fragmentation patterns are identical to those of the deprotonated vanillin. The confirmation of structure of m/z 151.0397 ion is done by comparison of the retention time and HCDMS/MS data with commercial vanillin, which indeed shows the same with the unknown peak from post Fenton reaction extract, shown in Figure 4.9 upper panel. Hence, this unknown peak at rt 8.96 min in the trilignol degradation product mixture is identified as vanillin. There are another two clusters of chromatographic peaks shown in the Figure.4.9 lower panel that have the ions of 151.0397 at different retention time than standard vanillin. One of the possible explanations is that observation of ion 151.0397 could be the result of in-source fragmentation of the other post-pretreatment products.
Figure 4.9. Extract ion chromatogram (EIC) and HCD/MS/MS spectrum of vanillin (upper panel), EIC and HCD/MS/MS spectrum of unknown m/z 151.0397 (lower panel)
Similarly, another two unknown peaks with retention time 1.73min and 10.54min are identified as vanillic acid and dimeric G(β-O-4)G (Figure 4.10 and 4.11) by comparing the retention times as well as HCD/MS/MS spectra with standard compounds. As showed in Figure 4.11, there are two clusters peak around 10.43 and 12.45 min that have the same exact m/z value as well as MS/MS spectrum, which indicates these are likely to be the isomers of vanillin that have very similar structure with vanillin. Similar observation happens to dimeric G(β-O-4)G, a pair of peaks at 12.18 min is not able to distinguish from pair of peaks at 10.53 min by the HCD/MS/MS experiment other than the different in their retention times. It could be result of being different diastereomers of dimer. The putative structures of other three unknown molecules in the trilignol degradation reaction mixture are proposed according to the exact mass as well as the fragmentation pattern (shown in Figure 4.12). Unfortunately, the lack of authentic compounds makes the identification and confirmation of these unknowns impossible.
Figure 4.10. EIC and HCD/MS/MS spectrum of vanillic acid (upper panel), EIC and HCD/MS/MS spectrum of m/z 167.0346 (lower panel)
Figure 4.11. EIC and HCD/MS/MS spectrum of synthetic G(β-O-4)G (upper panel), EIC and HCD/MS/MS spectrum of m/z 375.1439 (lower panel)
Figure 4.12. HCD/MS/MS spectra of three unknown features with putative structures.
4.4 Conclusion

Previous studies have shown that solution phase Fenton chemistry can play a positive role in increasing the accessibility of cellulose in biomass. However, the total lignin content in the biomass remain unchanged after solution phase Fenton pretreatment\textsuperscript{53}. Monitoring the reactions of the solution phase Fenton chemistry on the synthetic trimeric lignin model compound provides insight into understanding the degradation process of Fenton reagent towards lignin compounds. The time course measurement of trilignol shows that the trimer is degraded or modified in the solution phase Fenton condition within a few hours. The study also shows that the degradation is not a first order reaction. The concentration of starting trimer drops rapidly in the first 5min then undergoes at a slower degradation rate. A list of features was generated by using MZmine2 for data preprocessing. With high-resolution mass spectrometry, the possible lignin-like features could be screened out from the other interference ions according to the accurate mass and tandem mass spectrometry. Upon comparison with three authentic compounds, the production of vanillin, vanillic acid and G(β-O-4)G dimeric compound in the Fenton pretreatment reaction are successfully confirmed. While no significant lignin content decrease upon the Fenton pretreatment biomass (switchgrass and miscanthus), the Fenton reaction upon lignin model compounds shows opposite results. One possible explanation is that the accessibility of lignin reactive site toward hydroxyl radical. The hydroxyl radical is a strong oxidative radical with short lifetime and tend to randomly attack reactive sites in close proximity\textsuperscript{49}. During the study of the solution phase Fenton chemistry upon synthetic lignin model compound, the reaction was done in a homogeneous system. The trimeric compound was completely dissolved in the
presence of ethanol compared to the heterogeneous conditions found in ground biomass pretreatment. Moreover, due to the complexity of cell wall structure, lignin has complex cross-linkages with hemicellulose and other molecules. The complex lignin structure could largely reduce the available lignin reactive sites to free hydroxyl radicals. This study has successfully developed the LC-MS/MS analytical techniques for the trimeric lignin model compound degradation products. The study also provides an insightful understanding of the Fenton chemistry towards the lignin model compound which could be extended to natural lignin. This study raises up the potentials for improving reaction conditions, which is required for treating natural biomass and the possibility of complete removal of lignin with Fenton chemistry.
Chapter 5

Conclusion

The focus of this dissertation is to develop mass spectral analytical methods in characterization of lignin model compounds, design synthetic route for novel lignin model compounds as well as study biomass pretreatment using state of the art high-resolution mass spectrometry. Lignin has drawn increasing attention in the area of biomass-to-biofuel conversion as it negatively affects the utilization of enclosed cellulose in the biomass. Un-locking the aromatics in lignin provides a potential renewable replacement for the traditional chemical precursors derived from petro-chemistry. A better understanding of lignins’ physical and chemical properties will contribute in exploring potentials to upgrade the commercial value of lignin, which requires complete elucidation of lignin structure. However, lignin analysis remains quite challenging, as no analytical methods have been able to completely depict the primary structures of lignin today. As mass spectrometry and tandem mass spectrometry have been successfully in elucidation of primary structure of proteins and peptides, it becomes a very attractive research area for the application of mass spectrometry in lignin characterization and structural elucidation.

One of the most important factors when applying mass spectrometry in lignin chemistry is to evaluate the ionization response of lignin compounds. Electron ionization (EI) coupled with gas chromatogram (GC)-mass spectrometry (MS) has been successfully applied in characterizing monomeric composition of biomass with the aid of prior chemical degradation. However, electron ionization (EI) is not compatible with non-volatile analytes. Interpretation of complex EI mass spectrum remains very
challenging, thus little structural information can be easily obtained for high molecular weight of lignin oligomers with traditional EI ionization technique. Soft-ionization techniques, such as electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI), have shown successful application in large biopolymers. Efforts have been investigated in characterizing lignin compounds with several soft-ionization techniques coupled with mass spectrometry, such as ESI, MALDI, Atmospheric Pressure Chemical Ionization (APCI) and Atmospheric Pressure Photoionization (APPI). Among these, electrospray ionization (ESI) is widely used since it requires less rigorous sample preparation and is easily compatible with separation techniques, such as high-performance liquid chromatography (HPLC). Electrospray ionization allows easy selection of positive ion or negative ion modes, which is mainly dependent on the chemical functional groups of the analytes. In an attempt to apply ESI in lignin study, negative ionization mode in ESI through simple deprotonation is commonly used because of the presence of weak acidic phenol functional group. Due to the structural diversity of lignin, negative ion mode with traditional deprotonation techniques may not be suitable for the lignin oligomers that don’t have free phenolic hydroxyl group. A few studies have observed that positive ion mode in ESI through adduct formation gives better ionization response than negative ion mode via simple deprotonation20-21,32.

The second chapter of this dissertation focuses on systematic comparison of ionization response with known structures of lignin trimeric model compounds under different ESI conditions. The preliminary mass spectral characterization of dehydrogenation polymers (DHPs) synthesized in two different reaction conditions shows the significant structural difference in the product profile. The DHP synthesized in
the presence of CTAS (DHP-CT) contained high abundant of α, β-diaryl ether linkage. The ESI ionization response of the DHP-CT showed distinct difference between negative ion ESI via simple deprotonation and positive ion mode ESI via the formation of ammonium cation adducts. Higher molecular weight oligomers, number of degree of polymerization to 10, were observed in positive ion mode of α, β-diaryl ether linkage rich DHP, while same high molecular weight ions were absent in the negative ion mode with simple deprotonation. Monomeric and dimeric ions were the most abundant in negative ion mode. In the sharp contrast to DHP-CT, DHPs, synthesized from continuous addition in aqueous buffer (DHP-ZT), gave similar molecular weight distribution in both negative ion mode and positive ion mode. This interesting observation motived a further study for a deeper understanding of the relationship between the ionization conditions and the structures of lignin. This ionization study might provide insights for selecting ionization conditions when encountering lignin compounds. The first hypothesis of this dissertation is that the ionization response of lignin is highly related to the specific chemical linkage types of lignin compounds as well as the solution chemistry of electrospray ionization. This hypothesis was tested by a systematic comparison of three trilignols containing unique chemical bonding type under different ESI conditions. The results of negative ion mode ESI with deprotonation [M-H]- and chloride adduction formation [M+Cl]- showed that the trimer containing α, β-diaryl ether linkage suffered significantly from in-source fragmentation. When forming [M-H], α, β-diaryl ether trimer resulted in poor ion response compared with other trimers containing typical β-O-4 linkage. Through detail comparison of the structure of each trimer, it was proposed that the trimer with α, β-diaryl ether linkage requires less fragmentation energy because it
contains phenolic group at α-position as a much better leaving group compared with hydroxyl group at α-position in the other two trimers. This in-source fragmentation is especially activated when [M-H]⁻ generated by deprotonation at the phenolic site. The deprotonation activates the charge-driven fragmentation pathway. The result of chlorine adduct formation [M+Cl]⁻ has supported this hypothesis by obtaining negative ions without the removal of proton at phenolic site.

As the preliminary data showed, ammonium adduction formation gave overall good response across two different DHPs. Quantitative comparison of a series of different cation adduct formations was conducted with trimeric model compounds. Results from these experiments showed that Li⁺ and Na⁺ were the two cations that gave best ion response across all three trimers. Compared with negative ion mode, no in-source fragment ions were observed in the positive ion mode. This result was consistent with previous observations. Interestingly, the trimer, containing α, β-diaryl ether linkage, gave an overall much better ion response compared with other two trimers across all different cation condition in the positive ion mode. It is proposed that α, β-diaryl ether linkage trimer contains no cyclic linkage type and has higher degree of freedom to rotate in order to coordinate with cations compared with other two trimers with cyclic linkage present.

Alkali metal adduct ions in collision induced dissociation (CID) are known to give little or no structural information as either alkali metal cation is fragmented off immediately or no fragment ions are generated during CID process. In contrast to traditional CID observation towards alkali metal adduct ion, useful fragmentation ions during the tandem mass spectrometric experiment using higher-energy collision
dissociation (HCD) techniques were successfully obtained. Fragmenting alkali metal adduct ions, such as [M+Li]^+ and [M+Na]^+, generally required higher collision energy than that of [M+NH_4]^+ or [M+H]^+. The HCD/MS/MS study allows the investigation of: 1) whether there is difference in fragmentation pattern of lignin compounds between different alkali metal cations; 2) whether useful and characteristic structural information could be obtained based on the fragmentation of alkali metal adduct cations. This study extended the understanding regarding the behavior of these alkali metal adduct ions in the higher-collision dissociation cells and has shown the promising application in studying the primary structure of lignin with this method.

The first project has deepened the knowledge of ESI ionization behavior of lignin model compounds under different condition as well as explored the potential application in sequencing lignin structures. This study provides a possible solution for natural lignin analysis in the future studies by adduct formation in positive ion mode. As the data shows good ionization response regardless of the structures of lignin compounds in the positive ion mode. As mentioned previously, the increasing interest in structural elucidation of higher molecular weight lignin compounds. One of the challenges in direct characterization of these lignin oligomers is the poor solubility of lignin compounds. Generally, further derivatizations of lignin oligomers, such as acetylation, are used to improve the solubility. Almost all the derivatization methods targets the free alcoholic groups, including phenolic group. Such derivatization, e.g. acetylation, caps the protic alcohol groups, which closes the possibility to form the deprotonated anion when ionization of higher lignin oligomers. As a result, formation of alkali metal adduct
provides alternative ionization strategies in ionizing these lignin oligomers, which are not favored in protonation or deprotonation in the traditional ESI conditions.

Further application of this ionization method on the lignin model compounds containing other unique structures would extend the knowledge regarding the fragmentation pattern in existence of different chemical bonding types. It will be beneficial to generate a library which contains the useful information including the lignin structures and their fragmentation patterns under different ionization conditions. This kind of library will be beneficial for those who intend to study primary structures of unknown lignin compounds by using tandem mass spectrometry in the future.

In the early biomass pretreatment experiments with miscanthus and switch grass using Fenton chemistry, positive response in the cellulose accessibility is observed by comparing the biomass before and after pretreatment\textsuperscript{53}. However, these studies of lignin show that no difference in either the total lignin content of biomass or the relative monomeric composition before and after pretreatment. This question motivated the continuous study of the actual chemical effect of Fenton chemistry during the pretreatment step. In order to simplify the complexity of starting pretreatment target, a synthesized lignin model compound with known structure was used to unambiguously study the role of Fenton chemistry by tracking the well-defined starting lignin model compound instead of using extracted native lignin. The object was to design a new lignin model compound that has not been obtained through chemical synthetic methods. Because different functional groups in lignin have different reactivities towards pretreatment process, it is beneficial to synthesize such model compound that contains all these important functional groups, such as β-O-4 linkage, free phenolic alcoholic group
and α, β-unsaturated alcoholic side chain. The goal of this project is to synthesize such lignin model compound to study the pretreatment process. Aldol-type condensation reaction has previously been applied for synthesizing β-O-4 linkage\textsuperscript{44-45, 63}. However, the model compound designed in this dissertation contains α, β-unsaturated alcoholic side chain, which cannot be obtained using the previous published synthesis methods. This motivated the need to come up with new synthetic strategies. The goal of the second project was to design a total synthetic route of trimer that contains the functionalities mentioned above. Different protecting groups and reagents were investigated in order to successfully obtain compound of interest. This project has not only provided a successful synthetic method but also generated a library of analytical data, such as NMR, HCD/MS/MS and ESI ionization properties, to characterize the intermediate compounds, which will be valuable for structural elucidation for similar compounds. In the third chapter, the project not only successfully provided a route for synthesizing new model compound, but also provided the feasible synthetic strategy that was applicable for introducing structural varieties in the model compounds. Those structural diversities could be different monomeric sequence composition and the number of monomeric units within the oligomers in the future application. This method provided a potential extension in developing an automatic synthesis technique.

As the overall goal was to better understand the function of Fenton chemistry towards lignin based on the observation of the natural biomass discussed previously. The objective of the third project was to develop mass spectrometry-based analytical methods for studying Fenton chemistry on the synthetic trimeric model compound. The strong oxidant, hydroxyl radical, was the active reagent in the degradation of lignin, however,
due to the randomness and non-specificity of oxidation using hydroxyl radical, complex post pretreatment product profile was expected. In order to study the complex sample composition, mass spectrometry (MS) coupled with high-performance liquid chromatography (HPLC) was used to provide another dimension of separation along with mass spectrometric data of the complex pretreatment products. High-resolution accurate-mass mass data were obtained using Thermo Q Exactive Orbitrap mass spectrometer. The high resolution and accurate mass measurement allowed the assignment of elemental composition to the unknown ions and reduced the background interference ions during data processing. Due to the non-specificity of the Fenton pretreatment products, similar data mining strategy used in metabolic research was applied in chapter 4. Bioinformatics tools, such as XCMS and MZmine2, are capable to process the raw data and generate a list of features from raw LC-MS data. Each feature is a chromatographic peak containing information, such as mass-to-charge, retention time, peak area and etc. The features were detected according to the parameters set during data processing with MZmine2 (in this dissertation). This is a time-efficient way for untargeted analysis when no specific targeted information is available. By generating a list of features that meet the requirement, one can eliminate the numbers of ions, which are not likely to be the ions of interest. Because of the high-resolution accurate-mass mass spectrometric analysis in the study, twenty features were selected with a higher possibility to be related with lignin model compounds. The concentration change curve of starting trimer over 5hrs showed that the degradation/modification of trimer in the presence of Fenton reagent was not a linear chemical reaction. The trimer concentration was dropping rapidly in the first 5min and slowly went to completion in about 5hrs. The time course curves of the other features
were different from each other. For the trimeric unknown features, similar degradation trend with the starting trimer was observed. The time course measurements of a few monomeric and dimeric features showed that these possible pretreatment products were generated once the reaction began. The products were also subjected to further degradation in the Fenton reaction condition. Among all the unknown features, three features were confirmed by comparing the LC/MS as well as HCD/MS/MS spectra with authentic compounds. These three compounds were vanillin, vanillic acid and dimer G(β-O-4)G (rt = 10.54 min). The high-resolution accurate-mass data allowed the assignment of elemental composition to the unknown features. The corresponding HCD/MS/MS fragment ions of unknown features provided valuable information in proposing the possible chemical structures based on the formula and fragmentation patterns. However, for unambiguous structural assignment of these post-pretreated compounds, further confirmation of those putative structures is required through comparing with authentic compounds when available. In contrast to the lignin study of Fenton pretreatment of biomass, the results of this project have shown that solution phase Fenton chemistry can degrade a lignin model compound very effectively. One possible explanation for this observation was that the physical and chemical complexity of lignin model compound was much simpler than that of native lignin structure. In the native lignin, the existence of various chemical linkages, such as cyclic C-C condense bonding types and intermolecular linkages to hemicellulose, brings different levels of the recalcitrance during Fenton pretreatment. Moreover, hydroxyl radical is very strong oxidant that oxidizing chemicals in close proximity due to its short lifetime, while complex cell wall structures in the biomass decrease the exposure of lignin to the active hydroxyl radicals. The hypothesis of
this study is that the Fenton chemistry caused partial breakdown of lignin across whole biomass resulting in the increase of cellulose accessibility. This project has developed LC/MS/MS methods for analyzing lignin related compounds. This analytical method can be further applied to study other post-pretreatment products generated with different pretreatment methods.

My research has explored and developed mass spectrometry-based analytical methods in depicting the structures of lignin related compounds, especially for the ones that have molecular weight higher than trimers. The first project shows the potential ionization conditions that can provide good ionization response of a variety of lignin oligomers regardless of the primary structures using ESI-MS. The HCD/MS/MS experiments offered unique fragmentation properties in lignin structural elucidation method.

Continuing to synthesize various lignin model compounds with the new synthetic strategy can largely expand the library of lignin model compounds, which will be beneficial to systematically generate useful mass spectral database for structural elucidation of lignin. Moreover, it will be worthwhile to test whether the synthetic lignin model compounds have potential application in material science or chemical engineering area. This brings a promising strategy to increase the industrial value of lignin, as lignin is a green and sustainable bio-resource. The untargeted LC/MS method presented in this dissertation provided a new high-throughput analytical strategy in studying complex lignin pretreatment products. The overall objective of this dissertation is to provide deeper understanding and potential solutions to answer the challenging questions
regarding the structure of large lignin molecules or lignin oligomers with high-resolution mass spectrometry-based analytical methods.
References

1. EIA, Annual Energy Outlook. 2015.
21. Reale, S.; Attanasio, F.; Spreti, N.; De Angelis, F., Lignin Chemistry:
Biosynthetic Study and Structural Characterisation of Coniferyl Alcohol Oligomers
22. Reale, S.; Di Tullio, A.; Spreti, N.; De Angelis, F., Mass spectrometry in the
23. Metzger, J. O.; Bicke, C.; Faix, O.; Tuszynski, W.; Angermann, R.; Karas, M.;
Strupat, K., Matrix-Assisted Laser Desorption Mass Spectrometry of Lignins**.
24. Bocchini, P.; Galletti, G. C.; Seraglia, R.; Traldi, P.; Camarero, S.; Martinez, A. T.,
Matrix-assisted Laser Desorption/Ionization Mass Spectrometry of Natural and
25. De Angelis, F.; Fregonese, P.; Verì, F., Structural Investigation of Synthetic
Lignins by Matrix-assisted Laser Desorption/Ionization Time-of-flight Mass
Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry for Second-Generation
27. Yoshioka, K.; Ando, D.; Watanabe, T., A Comparative Study of Matrix- and
Nano-assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry of Isolated
28. Evtuguin, D. V.; Domingues, P.; Amado, F. L.; Neto, C. P.; Correia, A. J. F.,
Electrospray Ionization Mass Spectrometry as a Tool for Lignins Molecular Weight and
29. önnerud, H.; Palmblad, M.; Gellerstedt, G., Investigation of Lignin Oligomers
Goeminne, G.; Inzé, D.; Messens, E.; Ralph, J.; Boerjan, W., Mass Spectrometry-Based
Goeminne, G.; Inzé, D.; Messens, E.; Ralph, J.; Boerjan, W., Mass Spectrometry-
Based Fragmentation as an Identification Tool in Lignomics. *Analytical Chemistry* 2010,
82 (19), 8095-8105.
Amundson, L. M.; Narra, P.; Aqueel, M. S.; Parsell, T. H.; Abu-Omar, M. M.;
Kenttämäa, H. I., Characterization of model compounds of processed lignin and the
lignome by using atmospheric pressure ionization tandem mass spectrometry. *Fuel* 2012,
95 (0), 634-641.
33. Dean, J. C.; Burke, N. L.; Hopkins, J. R.; Redwine, J. G.; Ramachandran, P. V.;
McLuckey, S. A.; Zwier, T. S., UV Photofragmentation and IR Spectroscopy of Cold, G-
Type β-O-4 and β-β Dilignol–Alkali Metal Complexes: Structure and Linkage-
1917-1932.
50. Arantes, V.; Milagres, A. F.; Filley, T.; Goodell, B., Lignocellulosic polysaccharides and lignin degradation by wood decay fungi: the relevance of


VITA

Fan Huang was born in Pingnan, a small town in Fujian province, southeast of China on September 25, 1988. She stayed in Pingnan till graduated from middle school. She then moved to Fuzhou for high school in 2003. She was enrolled in undergraduate program in University of Science and Technology of China (USTC) in Hefei, Anhui, China, in 2006. Fan did her undergraduate research under the advising of Dr. Yangzhong Liu, where she got to know about mass spectrometry. After she graduated from USTC with a B.S. in Chemistry in 2010, she stayed in Dr. Yangzhong Liu’s group for another year. Fan was enrolled in Department of Chemistry, University of Kentucky in 2011. Fan was interested to study more about mass spectrometry and analytical chemistry, so she joined Dr. Bert C Lynn’s research group to pursue her Ph.D. in analytical chemistry in 2012. She is a member of the American Chemical Society (ACS) and the American Society for Mass Spectrometry (ASMS).

Publications:

1. F. Huang, B.C. Lynn, “Structural effects on the ionization response of lignin model compounds during electrospray ionization” (In process).

Presentations:

Appendices

Nuclear Magnetic Resonance (NMR) spectra for the compounds that were discussed in this dissertation. The NMR spectra were acquired with a Varian Model NOVA400 (400 MHz) spectrometer, otherwise specified.

All the compounds were run in the Acetone-d₆, acetone-d₆ signal was used as reference (δ 2.05ppm).

Figure S1. 1D proton NMR spectrum of coniferyl alcohol
Figure S2. 1D proton NMR spectrum of G(4-O-α)G(β-O-4)G (Varian NOVA 600)
Figure S3. 2D HSQC NMR spectrum of G(4-O-α)G(β-O-4)G (Varian NOVA 600)
Figure S4. 1D proton NMR spectrum of G(β-O-4)G(β-5)G
Figure S5. 1D proton NMR spectrum of G(β-O-4)G(β-β)G
Figure S6. 2D HSQC NMR spectrum of mixture of G(β-O-4)G(β-5)G and G(β-O-4)G(β-β)G
Figure S7. 1D proton NMR spectrum of Compound 4
Figure S8. 1D proton NMR spectrum of Compound 2
Figure S9. 1D proton NMR spectrum of Compound 3
Figure S10. 1D proton NMR spectrum of Compound 5
Figure S11. 1D proton NMR spectrum of Compound 6
Figure S12. 1D proton NMR spectrum of Compound 9 (Varian NOVA 600)
Figure S13. 2D HSQC NMR spectrum of Compound 9 (Varian NOVA 600)
Figure S14. 1D proton NMR spectrum of Compound 13
Figure S15. 2D HSQC NMR spectrum of Compound 13
Table S1. Deprotonated ion intensity comparison of three trimers
Table S2. Ion intensity comparison of three trimers with chloride ion adduct, [M+Cl]⁻ in negative ion mode ESI.
Table S3. a) Relative ionization response of G(4-O-α)G(β-O-4)G in different salt solution. b) Cation selectivity of G(4-O-α)G(β-O-4)G in solution of equal molar of NH₄Cl/LiCl/NaCl/KCl/RbCl.
Table S4.  a) Relative ionization response of G(β-O-4)G(β-5)G in different salt solution. b) Cation selectivity of G(β-O-4)G(β-5)G in solution of equal molar of NH₄Cl/LiCl/NaCl/KCl/RbCl.
Table S5. a) Relative ionization response of \( \text{G(\(\beta\)-O-4)G(\(\beta\)-\(\beta\))G} \) in different salt solution. b) Cation selectivity of \( \text{G(\(\beta\)-O-4)G(\(\beta\)-\(\beta\))G} \) in solution of equal molar of \( \text{NH}_4\text{Cl/LiCl/NaCl/KCl/RbCl} \).

Table S6. Positive ion mode ESI response comparison of three trilignols with the presence of different cations.
Figure S1. 1D proton NMR spectrum of coniferyl alcohol
Figure S2. 1D proton NMR spectrum of G(4-O-α)G(β-O-4)G
Figure S3. 2D HSQC NMR spectrum of G(4-O-α)G(β-O-4)G
Figure S4. 1D proton NMR spectrum of G(β-O-4)G(β-5)G
Figure S5. 1D proton NMR spectrum of G(β-O-4)G(β-β)G
Figure S6. 2D HSQC NMR spectrum of mixture of G(β-O-4)G(β-5)G and G(β-O-4)G(β-β)G
Figure S7. 2D HSQC NMR spectrum of mixture of G(β-O-4)G(β-5)G and G(β-O-4)G(β-β)G
Figure S8. 1D proton NMR spectrum of Compound 2
Figure S9. 1D proton NMR spectrum of Compound 3
Figure S10. 1D proton NMR spectrum of Compound 5
Figure S11. 1D proton NMR spectrum of Compound 6
Figure S12. 1D proton NMR spectrum of Compound 9
Figure S13. 2D HSQC NMR spectrum of Compound 9
Figure S14. 1D proton NMR spectrum of Compound 13
Figure S15. 2D HSQC NMR spectrum of Compound 13
Table S1. Deprotonated ion intensity comparison of three trimers

<table>
<thead>
<tr>
<th>Trimer</th>
<th>Average Intensity</th>
<th>STDEV</th>
<th>%RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>G(4-O-α)G(β-O-4)G</td>
<td>13908.39</td>
<td>5974.62</td>
<td>42.96</td>
</tr>
<tr>
<td>G(β-O-4)G(β-5)G</td>
<td>282557.90</td>
<td>21803.84</td>
<td>7.72</td>
</tr>
<tr>
<td>G(β-O-4)G(β-β)G</td>
<td>470547.78</td>
<td>60802.74</td>
<td>12.92</td>
</tr>
</tbody>
</table>
Table S2. Ion intensity comparison of three trimers with chloride ion adduct, [M+Cl]− in negative ion mode ESI.

<table>
<thead>
<tr>
<th>Trimer</th>
<th>Average Intensity</th>
<th>STDEV</th>
<th>%RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>G(4-O-α)G(β-O-4)G</td>
<td>1914943.12</td>
<td>55030.00</td>
<td>2.87</td>
</tr>
<tr>
<td>G(β-O-4)G(β-5)G</td>
<td>1147487.21</td>
<td>46184.23</td>
<td>4.02</td>
</tr>
<tr>
<td>G(β-O-4)G(β-β)G</td>
<td>2489207.77</td>
<td>566208.40</td>
<td>22.75</td>
</tr>
</tbody>
</table>
Table S3. a) Relative ionization response of G(4-O-α)G(β-O-4)G in different salt solution. b) Cation selectivity of G(4-O-α)G(β-O-4)G in solution of equal molar of NH₄Cl/LiCl/NaCl/KCl/RbCl.

a)

<table>
<thead>
<tr>
<th>Different Cations</th>
<th>Average Ion Response</th>
<th>STDEV</th>
<th>%RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₄</td>
<td>42.04</td>
<td>7.91</td>
<td>18.82</td>
</tr>
<tr>
<td>Li</td>
<td>93.04</td>
<td>6.79</td>
<td>7.30</td>
</tr>
<tr>
<td>Na</td>
<td>100.00</td>
<td>12.23</td>
<td>12.23</td>
</tr>
<tr>
<td>K</td>
<td>45.76</td>
<td>2.84</td>
<td>6.20</td>
</tr>
<tr>
<td>Rb</td>
<td>32.45</td>
<td>0.95</td>
<td>2.94</td>
</tr>
</tbody>
</table>

b)

<table>
<thead>
<tr>
<th>Different Cations</th>
<th>Average Ion Response</th>
<th>STDEV</th>
<th>%RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₄</td>
<td>19.46</td>
<td>0.83</td>
<td>4.26</td>
</tr>
<tr>
<td>Li</td>
<td>100.00</td>
<td>4.50</td>
<td>4.50</td>
</tr>
<tr>
<td>Na</td>
<td>86.92</td>
<td>4.49</td>
<td>5.16</td>
</tr>
<tr>
<td>K</td>
<td>41.94</td>
<td>1.93</td>
<td>4.59</td>
</tr>
<tr>
<td>Rb</td>
<td>26.01</td>
<td>0.84</td>
<td>3.22</td>
</tr>
</tbody>
</table>
Table S4.  

a) Relative ionization response of G(β-O-4)G(β-5)G in different salt solution.
b) Cation selectivity of G(β-O-4)G(β-5)G in solution of equal molar of NH₄Cl/LiCl/NaCl/KCl/RbCl.

<table>
<thead>
<tr>
<th>Different Cations</th>
<th>Average Ion Response</th>
<th>STDEV</th>
<th>%RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₄</td>
<td>94.00</td>
<td>2.35</td>
<td>2.50</td>
</tr>
<tr>
<td>Li</td>
<td>81.60</td>
<td>13.55</td>
<td>16.61</td>
</tr>
<tr>
<td>Na</td>
<td>100.00</td>
<td>5.27</td>
<td>5.27</td>
</tr>
<tr>
<td>K</td>
<td>89.00</td>
<td>4.61</td>
<td>5.18</td>
</tr>
<tr>
<td>Rb</td>
<td>63.27</td>
<td>1.28</td>
<td>2.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Different Cations</th>
<th>Average Ion Response</th>
<th>STDEV</th>
<th>%RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₄</td>
<td>43.72</td>
<td>5.98</td>
<td>13.68</td>
</tr>
<tr>
<td>Li</td>
<td>91.28</td>
<td>16.25</td>
<td>17.81</td>
</tr>
<tr>
<td>Na</td>
<td>100.00</td>
<td>15.68</td>
<td>15.68</td>
</tr>
<tr>
<td>K</td>
<td>80.93</td>
<td>10.11</td>
<td>12.49</td>
</tr>
<tr>
<td>Rb</td>
<td>59.70</td>
<td>7.03</td>
<td>11.77</td>
</tr>
</tbody>
</table>
Table S5. a) Relative ionization response of G(β-O-4)G(β-β)G in different salt solution. b) Cation selectivity of G(β-O-4)G(β-β)G in solution of equal molar of NH₄Cl/LiCl/NaCl/KCl/RbCl.

<table>
<thead>
<tr>
<th>Different Cations</th>
<th>Average Ion Response</th>
<th>STDEV</th>
<th>%RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₄</td>
<td>46.82</td>
<td>4.26</td>
<td>9.09</td>
</tr>
<tr>
<td>Li</td>
<td>93.46</td>
<td>4.09</td>
<td>4.37</td>
</tr>
<tr>
<td>Na</td>
<td>100.00</td>
<td>7.97</td>
<td>7.97</td>
</tr>
<tr>
<td>K</td>
<td>72.78</td>
<td>4.47</td>
<td>6.14</td>
</tr>
<tr>
<td>Rb</td>
<td>76.55</td>
<td>5.54</td>
<td>7.23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Different Cations</th>
<th>Average Ion Response</th>
<th>STDEV</th>
<th>%RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₄</td>
<td>29.21</td>
<td>2.80</td>
<td>9.59</td>
</tr>
<tr>
<td>Li</td>
<td>100.00</td>
<td>9.31</td>
<td>9.31</td>
</tr>
<tr>
<td>Na</td>
<td>96.80</td>
<td>9.25</td>
<td>9.55</td>
</tr>
<tr>
<td>K</td>
<td>55.69</td>
<td>4.52</td>
<td>8.12</td>
</tr>
<tr>
<td>Rb</td>
<td>33.00</td>
<td>2.25</td>
<td>6.81</td>
</tr>
</tbody>
</table>
Table S6. Positive ion mode ESI response comparison of three trilignols with the presence of different cations.

<table>
<thead>
<tr>
<th>Trimer</th>
<th>Cations</th>
<th>Average Ion response</th>
<th>STDEV</th>
<th>%RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>G(4-O-α)G(β-O-4)G</td>
<td>NH4</td>
<td>42.04</td>
<td>7.91</td>
<td>18.82</td>
</tr>
<tr>
<td></td>
<td>Li</td>
<td>93.04</td>
<td>6.79</td>
<td>7.30</td>
</tr>
<tr>
<td></td>
<td>Na</td>
<td>100.00</td>
<td>12.23</td>
<td>12.23</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>45.76</td>
<td>2.84</td>
<td>6.20</td>
</tr>
<tr>
<td></td>
<td>Rb</td>
<td>32.45</td>
<td>0.95</td>
<td>2.94</td>
</tr>
<tr>
<td>G(β-O-4)G(β-5)G</td>
<td>NH4</td>
<td>10.59</td>
<td>0.26</td>
<td>2.50</td>
</tr>
<tr>
<td></td>
<td>Li</td>
<td>9.20</td>
<td>1.53</td>
<td>16.61</td>
</tr>
<tr>
<td></td>
<td>Na</td>
<td>11.27</td>
<td>0.59</td>
<td>5.27</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>10.03</td>
<td>0.52</td>
<td>5.18</td>
</tr>
<tr>
<td></td>
<td>Rb</td>
<td>7.13</td>
<td>0.14</td>
<td>2.03</td>
</tr>
<tr>
<td>G(β-O-4)G(β-β)G</td>
<td>NH4</td>
<td>13.28</td>
<td>1.21</td>
<td>9.09</td>
</tr>
<tr>
<td></td>
<td>Li</td>
<td>26.52</td>
<td>1.16</td>
<td>4.37</td>
</tr>
<tr>
<td></td>
<td>Na</td>
<td>28.37</td>
<td>2.26</td>
<td>7.97</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>20.65</td>
<td>1.27</td>
<td>6.14</td>
</tr>
<tr>
<td></td>
<td>Rb</td>
<td>21.72</td>
<td>1.57</td>
<td>7.23</td>
</tr>
</tbody>
</table>