2013

Bio-mass for biomass: biological mass spectrometry techniques for biomass fast pyrolysis oils

Erica A. Dalluge
Iowa State University

Bio-mass for biomass:
Biological mass spectrometry techniques for biomass fast pyrolysis oils

by

Erica A. Dalluge

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Analytical Chemistry

Program of Study Committee:
Young-Jin Lee, Major Professor
Robert S. Houk
Emily Smith
Robert C. Brown
Edward Yu

Iowa State University
Ames, Iowa
2013

Dedicated to my parents, Scott and Tammy Smith,

for their unconditional love and support

TABLE OF CONTENTS

Page

NONMENCLATURE..... iv

ACKNOWLEDGEMENTS v

ABSTRACT vi

CHAPTER I INTRODUCTION 1

Biomass Fast Pyrolysis and its Characterization 1
Kinetics of Biomass Fast Pyrolysis 4
High Resolution Mass Spectrometry 6
Research Objective and Approach 8
Dissertation Organization 8

CHAPTER II PETROLEOMIC ANALYSIS OF BIO-OILS FROM THE FAST PYORLYSIS OF BIOMASS: LASER DESORPTION IONIZATON- LINEAR ION TRAP- ORBITRAP MASS SPECTROMETRY APPROACH 12

CHAPTER III BIO-OIL ANALYSIS USING NEGATIVE ELECTROSPRAY IONIZATION: COMPARATIVE STUDY OF HIGH RESOLUTION MASS SPECTROMETERS AND PHENOLIC VERSUS SUGARIC COMPOUNDS 43

CHAPTER IV STITCHING HRMS SPECTRA TOGETHER FOR A MORE COMPREHENSIVE PICTURE OF BIO-OILS 74

CHAPTER V REAL-TIME MONITORING OF BIOMASS FAST PYROLYSIS WITH μPYROLSIS-ATMOSPHERIC PRESSURE CHEMICAL IONIZATION TIME-OF-FLIGHT MASS SPECTROMETER 91

CHAPTER VI GENERAL CONCLUSIONS 127

Conclusions 126
Future Directions 127

APPENDIX INSIGHTS INTO BIO-OIL AGING USING A COMPREHENSIVE PETRLEOMIC CHARACTERIZATION WITH FTICR 128

NOMENCLATURE

HRMS	High Resolution Mass Spectrometry
FTICR	Fourier Transform Ion Cyclotron Resonance
MS	Mass Spectrometer
M/Z	Mass-to-charge
ESI	Electrospray Ionization
APPI	Atmospheric Pressure Photoionization
APCI	Atmospheric Pressure Chemical Ionization
µPy	Micropyrolysis
DBE	Double bond equivalency
GPC	Gel Permeation Chromatography
GC	Gas Chromatography
BIO-OIL	Biomass Pyrolysis Oils

ACKNOWLEDGEMENTS

I would like to thank my major professor, Young-Jin Lee, and my committee members (Emily Smith, Sam Houk, Robert Brown, and Ed Yu) for their guidance and support throughout the course of this research. I would also like to thank Theresa Windus for serving as a non-academic mentor throughout my career as a graduate student. Without the guidance and support of these faculty members, this research and my success would not have been possible.

In addition, I would also like to thank my fellow Lee Group colleagues and Brown Group members for their helpful discussions and suggestions during my research endeavors. These individuals, along with my friends, played a pivotal role in making my time at Iowa State University a wonderful experience.

Finally, thanks to my family for their encouraging words and reassuring confidence in my abilities, especially in times when I doubted myself. And to my husband, thank you for the unwithering patience, love, and respect you had for me during this adventure (and for the insightful and sometimes unwanted discussion about research at the dinner table).

ABSTRACT

Biomass fast pyrolysis oils, or bio-oils, are a promising renewable energy source to supplement or replace petroleum-based products and fuels. However, there is a current lack of understanding about the pyrolysis process which creates a bottleneck towards making biomass pyrolysis an economically feasible option. In order to address this bottleneck, this research focuses on developing high resolution mass spectrometry (HRMS) techniques to address biomass pyrolysis at the molecular-level.

The first attempt at analyzing bio-oils with HRMS employs laser desorption ionization and LTQ-Orbitrap MS to successful identify over 100 compounds. These compounds consist of 3-6 oxygens and have double-bond equivalents (DBE) of 9-17. A petroleomic analysis and comparison of the bio-oil to the low-mass components in hydrolytic lignin suggest that these compounds are dimers and trimers of depolymerized lignin. A wider variety of bio-oil compounds, specifically volatile and non-volatile compounds, could be characterized with electrospray ionization (ESI). Specifically, (-) ESI allows for the characterization of over 800 molecular compounds, of which about 40 of these were previously known in GC-MS. These compounds include cellulose- and hemicellulose-derived pyrolysis products as well as lignin-derived pyrolysis products.

A comparative study of three common HRMS was also performed to validate the methodology and to investigate differences in mass discrimination and resolution. This led to the development of a novel spectral stitching technique that combines datasets from different HRMS together. By stitching the datasets together inherent

instrument limitations (e.g. like mass discrimination and resolution) can be addressed. The resulting stitched mass spectrum gives rise to a more comprehensive picture of bio-oil.

Lastly, a pioneering technique that utilizes HRMS to monitor biomass fast pyrolysis in real-time has been developed. A fast-scanning time-of-flight mass spectrometer with a soft ionization source and a drop-in micropyrolyzer is used to provide insights into biomass pyrolysis that are not possible with traditional techniques. For example, metastable intermediates of cellulose pyrolysis could be identified and monitored with this novel approach. Also, fundamental pyrolysis studies, such as the effect of biomass shape and thickness, are possible with this technique due to the high sensitivity and time resolution of the time-of-flight mass spectrometer.

CHAPTER I

INTRODUCTION

A current global trend in research has been centered on developing renewable energy to supplement and eventually replace petroleum-based energy and nuclear energy. This demand for renewable energy arises from environmental and safety concerns with fossil fuels and nuclear energy. The depletion of fossil fuels has also driven research in this area due to the increased anxiety caused by the fear that fossil fuels are not a feasible long-term energy solution, especially with an exponential growth in global energy demand.[1]

Biomass Fast Pyrolysis and its Characterization

Fast pyrolysis of biomass is a promising technique for renewable energy that converts organic material (biomass) into bio-oil, char, and syngas by rapidly heating the biomass in the absence of oxygen to temperatures near 500 °C. The bio-oil that is obtained has a superficial physical resemblance to petroleum, but the chemical composition is quite different. Bio-oil contains a high amount of oxygenated and polar compounds due to the lignin and holocellulose components of biomass, which are largely absent in petroleum. Bio-oils also include an aqueous phase and a water-insoluble phase, both of which contain volatile and non-volatile compounds that make bio-oil a very complex mixture. Bio-oil can be catalytically upgraded to alkanes and aromatic compounds to produce biofuels. This catalytic upgrading process closely

resembles the refining of petroleum crude oil into gasoline and diesel fuel.[2] This means that petroleum refining infrastructure could be utilized to produce biofuels once the catalytic upgrading of bio-oils is optimized.

One of the significant bottlenecks in optimizing the catalytic upgrading of bio-oils is the lack of understanding of bio-oil components at the molecular level. Current characterization techniques rely heavily on bulk property measurements, such as pH, water content, acidity, density, viscosity, and heating values.[3] Structural insights have also been made with nuclear magnetic resonance (NMR), Fourier transform infrared (FTIR) spectroscopy, and with GC-MS data.[3-5] However, these techniques are limited to functional groups present in the mixture for NMR and FTIR studies, and to volatile compounds in GC-MS. These constraints limit the full understanding of the chemical structures of the bio-oils, particularly the non-volatile compounds at the molecular level.[3]

Several mass spectrometric analyses have been performed for bio-oil produced from fast pyrolysis of biomass. GC-MS has been successfully used for the analysis of volatile organic components for many decades.[6] However, appropriate tools have been lacking for the analysis of non-volatile oil compounds. Electrospray ionization (ESI) has been used in petroleum crude oil analysis for soft ionization of nonvolatile macromolecules,[7] particularly for its polar components. Based on this work, it is expected to be beneficial for bio-oil analysis because of the high percentage of polar compounds present. Low-energy electron ionization (EI) with a direct-probe sample introduction has been suggested as a possible alternative;[8] however, it suffers from low

ion signals and molecular fragmentations. Its application is also limited to partially volatile and thermally stable compounds. Field ionization (FI) is a valid approach for ionization of non-volatile compounds and has been used to study several biopolymers with a micropyrolyzer.[9, 10] Molecular beam MS has been used for the direct analysis of pyrolysis products with rapid quenching of the pyrolyzates, molecular-beam sampling and low-energy EI.[6] Single- or multi-photon ionization has also been applied for the analysis of biomass pyrolysis products.[11, 12] According to a study using a molecular-beam-EI quadrupole mass spectrometer on the pyrolysis products of various types of lignin by Evans and co-workers, molecular composition varies by both biomass material and chemical process.[6, 13]

Matrix-assisted laser desorption ionization (MALDI) or laser desorption ionization (LDI) has been used for molecular weight distributions of non-volatile compounds in bio-oils.[9] However, aggregate formation during the analysis can be problematic. Laser-induced aggregation has been investigated in LDI of coal asphaltenes by Hortal and co-workers.[14] At low laser power (15 μJ) and low sample concentration (2 mg/mL), a relatively narrow molecular-weight distribution centered at *m/z*~300 was observed. As the laser power increases, another molecular-weight distribution started to appear centered at *m/z* ~ 600, extending to *m/z* well above 1000. This high molecular-weight distribution is significant especially at higher sample concentration; this is consistent with the hypothesis that the aggregation reaction occurs in the laser plume. While decoupling desorption and ionization using a two photon laser system might be the best approach to solve laser induced aggregation ,[15] it is not easily adaptable or

commercially available. However, a significant limitation in these studies arises from the mass resolution of the mass analyzers used. Because of the limited mass resolutions adopted, most of the applications for studying bio-oil have simply shown molecular-weight distributions and have limited detailed chemical information available.[4-8]

Kinetics of Biomass Fast Pyrolysis

Studying the kinetics of biomass pyrolysis and the mechanisms involved is important to be able to control the end products of pyrolysis and to make biomass fast pyrolysis more competitive with petroleum based-products. This is why much attention has been spent on studying pyrolysis kinetics. Currently, most of the kinetic measurements have been made by thermogravimetric analysis (TGA) and μPy-GC/MS.[16-18] Unfortunately, TGA suffers from low heating rates that do not accurately resemble fast pyrolysis. This technique has only focused on obtaining a global or semi-global mechanism for specific biopolymers of biomass. This is possible because at lower heating rates each biopolymer (cellulose, hemicellulose, and lignin) has a unique decomposition temperature range.[19] However, at higher heating rates like in fast pyrolysis, the decomposition of the biopolymers is less distinguishable making TGA a less effective tool for studying the kinetics and mechanisms of fast pyrolysis.

There also has been much ambiguity in the reported results obtained from global mechanisms to explain kinetics. Experiments done on cellulose, the most simple and abundant biopolymer of biomass, have shown conflicting results in the calculated kinetic parameters.[17] Most of these experiments lump together vastly different chemical

compounds; therefore, they cannot explain the different degradation mechanisms that can occur and consequently affect the overall kinetics.

Research has been conducted to address the lack of non-global mechanisms in cellulose pyrolysis. Lédé and coworkers developed an innovative technique that utilizes a focusing 5kW xenon lamp to conduct radiant flash pyrolysis.[20] This system allows for a controlled heating time (down to 10ms) and for rapid quenching of gas and liquid intermediates. One limitation of this technique is that the exact pyrolysis temperature is unknown. In addition, the analysis is limited to an ex-situ analysis with GC-MS and LC-MS. This means that the liquid intermediates being analyzed may be different than those that escape from the reactor.

Another novel technique adapted by Dauenhauer and coworkers utilizes high-speed photography to capture images of cellulose pyrolysis.[21, 22] This allows for the dynamic nature of fast pyrolysis to be studied and confirmed the presence of a liquid intermediate known as molten biomass. This liquid intermediate/molten biomass has been previously suggested as active cellulose.[23] The presence of pyrolyzates being ejected as aerosols was also confirmed with high-speed photography.[22] Silylation followed by GC/MS of these aerosols indicate that they are mostly cellobiosan dimers that are formed through the direct decomposition of cellulose and then transported via aerosol ejection. Dauenhauer and coworkers also developed a thin-film cellulose pyrolysis technique that eliminated the conduction effects that larger particles (> 10 μm at 500 °C) have during pyrolysis.[17] This allowed for isothermal kinetics to be explored

and for an appropriate surrogate compound (α-cyclodextrin) to be identified for theoretical studies of cellulose pyrolysis.

Fundamental studies, like Dauenhauer and Lédé, have shown how complex the pyrolysis process is. The presence of a liquid intermediate suggests that pyrolysis is not only controlled by kinetics, but also dictated by dynamics of the liquid phase and its intermediates along with gas-phase reactions. Therefore, pyrolysis kinetics cannot be represented with global mechanisms or with simple one-component kinetic models. In order to construct a more encompassing mechanism, there needs to be more experimental data that addresses the liquid phase and its intermediates.[1] However, characterization of the liquid phase intermediates has largely been avoided due to the lack of analytical tools that can study the short-lived and complex intermediate species. Being able to analyze these intermediates would provide valuable information to develop more elaborate mechanisms of biomass fast pyrolysis and eventually allow for molecular-level kinetic information to be obtained.

High Resolution Mass Spectrometry

Recent developments in Fourier transform MS have enabled high-resolution mass spectrometric analysis of biological macromolecules using modern ionization techniques.[24] This allows for complex mixtures, like petroleum crude oil and bio-oil, to be sufficiently resolved to assign accurate masses and unique chemical compositions. Specifically, Fourier transform ion cyclotron resonance (FTICR) mass spectrometers allow for mass resolutions that exceed 200,000. With this advancement in resolving

capability, a renaissance of petroleum crude oil analysis has occurred that has been named "petroleomics" by Rodgers and Marshall.[25] Using an FTICR, Marshall's group could identify over 20,000 compounds in petroleum oils and developed systematic ways of analyzing and understanding these complex samples.[25] This approach was effectively used to compare molecular details of petroleum oils from different sources.[26] HRMS analysis of crude oils has been almost exclusively performed with Fourier transform ion cyclotron resonance (FTICR) mass spectrometers because of its superior mass resolving power that is needed for the enormous complexity of crude oils.

However, the use of only one type of mass spectrometer cannot avoid the instrument bias involved in mass spectrometric measurements. For example, a sample with a high oxygen content for low mass components may appear to be less oxygenated if the instrument has a significant mass discrimination in the low mass ion range. This could easily happen in the FTICR if one ignores the "time-of-fight effect" also known as ion-flight time.[4]

Recently, Pomerantz and co-workers proposed the use of an orbitrap mass analyzer for routine fingerprinting of major crude oil components and successfully demonstrated the similarity in heteroatom class distributions and double bond equivalent (DBE) distributions between FTICR and orbitrap data.[27] While this study suggests that an orbitrap mass analyzer may provide sufficient mass resolution for major crude oil compounds, other differences between the two instrumentations (such as mass discrimination) were not investigated. The operation principles between FTICR and orbitrap are quite different from each other, not only in mass analyzer but also in

ionization source design; therefore, a comparison is expected to reveal any instrument-dependent bias in the data.

Research Objective and Approach

This project aims at developing novel HRMS techniques to advance the current knowledge of biomass fast pyrolysis. Various high-resolution mass spectrometers and different ionization techniques were explored to characterize bio-oil in ways that traditional analytical techniques cannot. For example, this includes characterization of non-volatile and/or thermally unstable compounds which are not amenable to GC-MS analysis. Being able to analyze both volatile and non-volatile compounds allows for a more complete picture of bio-oil to be obtained. Also, a molecular-level characterization of bio-oil can be a valuable tool for biomass selection, process optimization, and downstream refining. This project also takes the HRMS methodology a step further by utilizing micropyrolysis-HRMS to study fundamentals of biomass pyrolysis in a unique and novel way. This is accomplished by using a fast-scanning time-of-flight MS, coupled to a micropyrolyzer, to monitor and analyze biomass fast pyrolysis in real-time.

Dissertation Organization

This dissertation is organized into six chapters with the first chapter (currently being read) encompassing a broad introduction. The second chapter utilizes orbitrap, a HRMS, with laser desorption ionization (LDI) to examine the bio-oil products that are derived from lignin. The third chapter utilizes a different ionization technique, negative

electrospray ionization (ESI), to examine the holocellulose pyrolysis products of biomass pyrolysis. This chapter also uses three different HRMS to ensure the profile of bio-oil is not being compromised with instrument bias. Chapter four expands on the potential for instrument bias, specifically low or high mass discrimination, and provides a novel data stitching technique to combat mass discrimination that is observed in orbitrap and FTICR with other ionization techniques besides negative ESI. Chapter five is about another novel HRMS technique that utilizes a µPy and TOF MS to monitor and analyze biomass fast pyrolysis in real-time. General conclusions and future directions are discussed in chapter six.

References

1. M. S. Mettler, D. G. Vlachos and P. J. Dauenhauer, *Energy & Environmental Science*, 2012, **5**, 7797-7809.
2. T. Marker, *Opportunities for biorenewables in oil refineries*, UOP LLC, 2005.
3. G. Gellerstedt, J. Li, I. Eide, M. Kleinert and T. Barth, *Energy Fuels*, 2008, **22**, 4240.
4. M. Kleinert and T. Barth, *Energy Fuels*, 2008, **22**, 1371.
5. A. Oasmaa and S. Czernik, *Energy Fuels*, 1999, **13**, 914.
6. G. Jiang, D. J. Nowakowski and A. V. Bridgwater, *Energy Fuels*, 2010, **24**, 4470.
7. S. Kim, R. P. Rodgers, G. T. Blakney, C. L. Hendrickson and A. G. Marshall, *J. Am. Soc. Mass Spectrom.*, 2009, **20**, 263.
8. F. Xu, Y. Xu, H. Yin, X. Zhu and Q. Guo, *Energy Fuels*, 2009, **23**, 1775.
9. R. Bayerbach, V. D. Nguyen, U. Schurr and D. Meier, *J. Anal. Appl. Pyrolysis*, 2006, **77**, 95.

10. R. J. Evans, T. A. Milne and M. N. Soltys, *J. Anal. Appl. Pyrolysis*, 1986, **9**, 207.

11. A. L. Brown, D. C. Dayton, M. R. Nimlos and J. W. Daily, *Chemosphere*, 2001, **42**, 663.

12. E. R. E. van der Hage, J. J. Boon, R. J. J. M. Steenvoorden and T. L. Weeding, *Anal. Chem.*, 1994, **66**, 543.

13. R. J. Evans and T. A. Milne, *Energy Fuels*, 1987, **1**, 123.

14. A. R. Hortal, P. Hurtado, B. Martinez-Haya and O. C. Mullins, *Energy Fuels*, 2007, **21**, 2863.

15. A. E. Pomerantz, M. R. Hammond, A. L. Morrow, O. C. Mullins and R. N. Zare, *Energy Fuels*, 2009, **23**, 1162.

16. J. E. White, W. J. Catallo and B. L. Legendre, *Journal of Analytical and Applied Pyrolysis*, 2011, **91**, 1-33.

17. M. S. Mettler, S. H. Mushrif, A. D. Paulsen, A. D. Javadekar, D. G. Vlachos and P. J. Dauenhauer, *Energy & Environmental Science*, 2012, **5**, 5414-5424.

18. M.-K. Bahng, C. Mukarakate, D. J. Robichaud and M. R. Nimlos, *Analytica Chimica Acta*, 2009, **651**, 117-138.

19. H. Yang, R. Yan, H. Chen, D. H. Lee and C. Zheng, *Fuel*, 2007, **86**, 1781-1788.

20. O. Boutin, M. Ferrer and J. Lédé, *Journal of Analytical and Applied Pyrolysis*, 1998, **47**, 13-31.

21. P. J. Dauenhauer, J. L. Colby, C. M. Balonek, W. J. Suszynski and L. D. Schmidt, *Green Chemistry*, 2009, **11**, 1555-1561.

22. A. R. Teixeira, K. G. Mooney, J. S. Kruger, C. L. Williams, W. J. Suszynski, L. D. Schmidt, D. P. Schmidt and P. J. Dauenhauer, *Energy & Environmental Science*, 2011, **4**, 4306-4321.

23. J. Lédé, *Journal of Analytical and Applied Pyrolysis*, 2012, **94**, 17-32.

24. T. M. Schaub, C. L. Hendrickson, S. Horning, J. P. Quinn, M. W. Senko and A. G. Marshall, *Anal. Chem.*, 2008, **80**, 3985.

25. A. Marshall and R. P. Rodgers, *Acc. Chem. Res.*, 2004, **37**, 53.

26. C. A. Hughey, C. L. Hendrickson, R. P. Rodgers, A. G. Marshall and K. Qian, *Anal. Chem.*, 2001, **73**, 4676.

27. A. Pomerantz, O. Mullins, G. Paul and M. Sanders, *Energy Fuels*, 2011, **25**, 3077.

CHAPTER II

PETROLEOMIC ANALYSIS OF BIO-OILS FROM THE FAST PYROLYSIS OF BIOMASS: LASER DESORPTION IONIZATION-LINEAR ION TRAP-ORBITRAP MASS SPECTROMETRY APPROACH

A paper published in *Energy and Fuels*

Erica A. Smith and Young Jin Lee*
Ames Laboratory, U. S. Department of Energy, Ames, Iowa and
Department of Chemistry, Iowa State University, Ames, Iowa 50011

Abstract

Fast pyrolysis of biomass produces bio-oils that can be upgraded into biofuels. In spite of similar physical properties to petroleum, the chemical properties of bio-oils are quite different and their chemical compositions, particularly those of non-volatile compounds, are not well known. Here, we report the first time attempt at analyzing bio-oils using high resolution mass spectrometry, which employed laser desorption ionization-linear ion trap-Orbitrap mass spectrometry. Besides a few limitations, we could determine chemical compositions for over 100 molecular compounds in a bio-oil sample produced from the pyrolysis of a loblolly pine tree. These compounds consist of 3-6 oxygens and 9-17 double bond equivalents (DBE). Among those, O_4 compounds with DBE of 9-13 were most abundant. Unlike petroleum oils, lack of near-by molecules within ±2 Da mass windows for major components enabled clear isolation of precursor ions for subsequent MS/MS structural investigations. Petroleomic analysis and comparison with low mass components in hydrolytic lignin suggest they are dimers and trimers of de-polymerized lignin.

Introduction

Among the many pathways for converting biomass into biofuels, the production of bio-oils via fast pyrolysis followed by catalytic upgrading to alkanes and aromatic compounds most closely resembles the refining of petroleum into gasoline and diesel fuel[1]. The bio-oil obtained by rapidly heating biomass in the absence of oxygen to temperatures near 500 °C has a superficial physical resemblance to petroleum, but the chemical composition is quite different. Bio-oil contains a high amount of oxygen due to the lignin and cellulose components of the biomass and polar compounds, which are largely absent in petroleum. Bio-oils include an aqueous phase and a water-insoluble phase, both of which contain non-volatile compounds that are difficult to detect by GC-MS (the conventional tool for bio-oil analysis).

One of the significant bottlenecks in optimizing the catalytic upgrading of bio-oil is the lack of molecular level understanding of bio-oil components. Current characterization techniques rely heavily on bulk property measurements such as pH, water content, acidity, density, viscosity, and heating values.[2] Great details of structural insights have been made with NMR and FT-IR and through GC-MS data.[3-5] However, the understanding is still limited to functional groups present in the mixture in NMR and FT-IR studies or volatile compounds in GC-MS. This limits the full understanding of the chemical structures of the bio-oils, particularly non-volatile compounds at the molecular level.[3]

Several mass spectrometric analyses have been performed for bio-oils produced from fast pyrolysis of biomass. GC-MS has been successfully utilized for the analysis of

volatile organic components for many decades.[6] However, appropriate tools have been lacking for the analysis of non-volatile oil compounds. Electrospray ionization allows soft ionization of non-volatile macromolecules and has been utilized in petroleum oil analysis[7], particularly for its polar components. It is expected to be beneficial for bio-oil analysis[3] because of the high percentage of polar compounds present. Low energy Electron Ionization (EI) with direct-probe sample introduction has been suggested as a possible alternative[8]; however, it suffers from low ion signals and EI fragmentations. Its application is also limited to partially volatile and thermally stable compounds. Field ionization (FI) is a valid approach for ionization of non-volatile compounds and has been used to study several biopolymers with an analytical pyrolyzer.[9, 10] Molecular beam mass spectrometry has been utilized for the direct analysis of pyrolysis products through rapid quenching with molecular-beam sampling and low energy electron ionization.[6] Single photon or multi-photon ionization has also been applied for the analysis of biomass pyrolysis products.[11, 12]

Matrix-assisted laser desorption ionization (MALDI) or laser desorption ionization (LDI) has been utilized for molecular weight distributions of non-volatile compounds in bio-oils along with size exclusion chromatography (SEC).[9] However, aggregate formation during the analysis could be problematic. Laser induced aggregation has been investigated in LDI of coal asphaltenes by Hortal and coworkers.[13] At low laser power (15 μJ) and low sample concentration (2 mg/ml) a relatively narrow molecular weight distribution centered at m/z ~ 300 was observed. As the laser power increases, another molecular weight distribution started to appear centered at m/z ~ 600

extending to m/z well above 1000. This high molecular weight distribution is significant especially with continuous extraction, compared to pulsed extraction, and at higher sample concentration; this is consistent with the hypothesis that the aggregation reaction occurs in the laser plume. While decoupling desorption and ionization using two photon laser system might be the best approach to solve the problem,[14] it is not easily adaptable or commercially available.

Another significant limitation comes from the mass resolution of the analyzers. Because of the limited resolutions adopted, most of the applications simply showed molecular weight distributions without detailed chemical information.[4-8] Double focusing mass spectrometers used to be the most popular high resolution mass spectrometers, but their application was mostly limited to small organic compounds due to the limitation in sample introduction. A remarkable exception is coupling a double focusing mass spectrometer with pyrolysis-field ionization to characterize organic matter in forest soils and lignin by Hempfling and Schulten.[15] They reported elemental compositions of twenty-five pyrolysis products of lignin and suggested possible dimeric structures. Recent developments in Fourier transform mass spectrometry enabled high resolution mass spectrometric analysis of biological macromolecules utilizing modern ionization techniques.[16] Specifically, Fourier transform ion cyclotron resonance mass spectrometers allow mass resolutions exceeding 200,000. With this advancement, along with the need to analyze complex asphaltenes in the petroleum industry, a renaissance of petroleum oil analysis has occurred, named ‘petroleomics’ by Rodgers and Marshall.[17] Marshall’s group could analyze over 20,000 compounds in petroleum oils and developed

systematic ways of analyzing and understanding these complex samples, such as Kendrick chart analysis and the carbon number vs. double bond equivalent (DBE) plot.[17] This approach was effectively utilized to compare molecular details of petroleum oils from different sources.[18]

According to a study using a molecular beam-electron ionization-quadrupole mass spectrometer on the pyrolysis products of various lignins by Evans and coworkers, molecular components could vary by both biomass material and chemical process.[6,19] Understanding and monitoring the crude bio-oils produced by fast pyrolysis could become an efficient tool to optimize the chemical process and improve the subsequent refinement, and to choose better biomass materials. Here reported is the first application of petroleomics to bio-oils produced by thermochemical conversion of biomass. Despite inherent limitations in the tools adapted for the current study (laser desorption ionization and Orbitrap mass analyzer) we could successfully demonstrate its application to fast pyrolysis of biomass and suggest some insight on the chemical process of biomass depolymerization.

Experimental

Chemicals

Hydrolytic lignin (Catalog #: 37-107-6), pyrene, and solvents were purchased from Sigma-Aldrich (St. Louis, MO) for the best available purity. The bio-oil samples were provided by Prof. Robert Brown at Iowa State University. In summary, the bio-oil used for this study was produced by fast pyrolysis of loblolly pine in a fluidized bed

pyrolyzer operated at 485 ºC. The sample was recovered from the third fraction in the bio-oil recovery system.[20] Further details on this system can be found elsewhere.[15] Bio-oils are dark-colored liquids that are chemically unstable and usually viscous. To slow chemical transformation, samples were refrigerated until the analysis.

Mass Spectrometric Data Acquisition

MALDI plates used for the analysis were thoroughly pre-washed through a deep-cleaning procedure suggested in the instruction manual provided by the manufacturer (Thermo Scientific). In short, the plate is sonicated first in 3 % ammonium hydroxide/ acetonitrile for 30 minutes, then in methanol. The samples were dissolved in three different HPLC grade solvents (2-propanol, methanol, and acetonitrile) from Sigma (St. Louis, MS) at a concentration of 1.0 mg mL^{-1}, and spotted on a MALDI plate in three increments of 0.5 µL, each being allowed to air dry between each increment. The spotted sample size is about 1 mm^2 which gives a spotted sample density of 1.5 $\mu g/mm^2$. These sample spots were analyzed with a linear ion trap-Orbitrap hybrid mass spectrometer (LTQ-Orbitrap Discovery; Thermo Scientific, San Jose, CA) using laser desorption ionization.[21] The Orbitrap analyzer was utilized for the mass spectral analysis with mass resolution of ~43,000 at m/z 272 and ~30,000 at m/z 400. The nitrogen laser (MNL 100; Lasertechnik Berlin, Berlin, Germany) used for the experiment has a wavelength of 337.7 nm, a maximum energy of 80 µJ/pulse and a maximum repetition rate of 60 Hz. The laser power used was between 5-35 µJ for the duration of the experiment. The actual laser power at the MALDI plate is expected to be 75% less than the initial power due to the power reduction caused by two neutral density

filters. Pyrene was used as an internal standard for the accurate mass calibration. The MALDI ion source pressure was maintained at 75-80 mTorr. Collision induced dissociation (CID) was performed in the LTQ with an isolation window of 1.8 Da (± 0.9 Da) and collision energy of 50 percent.

Results and Discussion

Mass Resolution of Orbitrap Analyzer

There have been two initial concerns in the approach we are taking. The first concern is that LDI may generate laser induced aggregation products. The other concern is that the Orbitrap mass analyzer, specifically the Discovery version we use, might not have sufficient mass resolution to resolve all the peaks in the mixture of bio-oil. The laser induced aggregation was investigated and minimized as will be discussed in the next section. The limited mass resolution is justified based on the fact that bio-oils are far less complex than petroleum oils. Most major peaks are at least two Dalton apart from each other, besides 13C isotopes, and there are only a few peaks, if any, within a Dalton of the mass window (this will be further discussed later; also see Figure 2A). To further confirm this hypothesis, the same sample was analyzed by a 7T FTICR (Solarix; Bruker, Billerica, MA) with a MALDI ion source at a Bruker service facility. Its mass resolution is ten times higher than that of Orbitrap Discovery: 450,000 at m/z 250 and 268,000 at m/z 400. According to this data, any two neighboring peaks are far apart from each other and expected to be resolved even with the lower mass resolution of Orbitrap Discovery (data not shown; MALDI FTICR data was not used for further

analysis because the laser power was not optimized). Hence, we concluded the bio-crude oil sample has much less complexity and Orbitrap mass analyzer provides sufficient resolution for the analysis. Mass resolution degrades rapidly as m/z increases in Fourier transform mass analyzers. This is why high mass resolution is particularly important in the analysis of petroleum asphaltene; however, bio-oils are mostly composed of low molecular weight components, with m/z 500 or lower, for which Orbitrap still seems to provide sufficient resolution.

Laser induced Aggregation in LDI

We investigated the laser induced aggregation of bio-oil to find optimal experimental conditions. We maintained low sample concentration and low laser power throughout the study, equivalent to or lower than those of Hortal's[9], to minimize the aggregation reactions. As shown in Figure 1A, LDI linear ion trap spectra of the bio-crude oil are composed of three major groups of peaks: low mass molecules of m/z 250-400 (Group I), medium mass molecules of m/z 400-550 (Group II), and high mass molecules of m/z 550 or higher (Group III). Group III is not very clear in Figure 1; we can only see some of them within the tail of Group II when we magnified its intensities, but it becomes apparent for a sample spotted with higher concentration that another band of peak distribution is clearly forming (data not shown). The absence of peaks below m/z ~ 250 in our spectra is due to the inability to see volatile compounds because they would evaporate from the MALDI plate before the analysis.

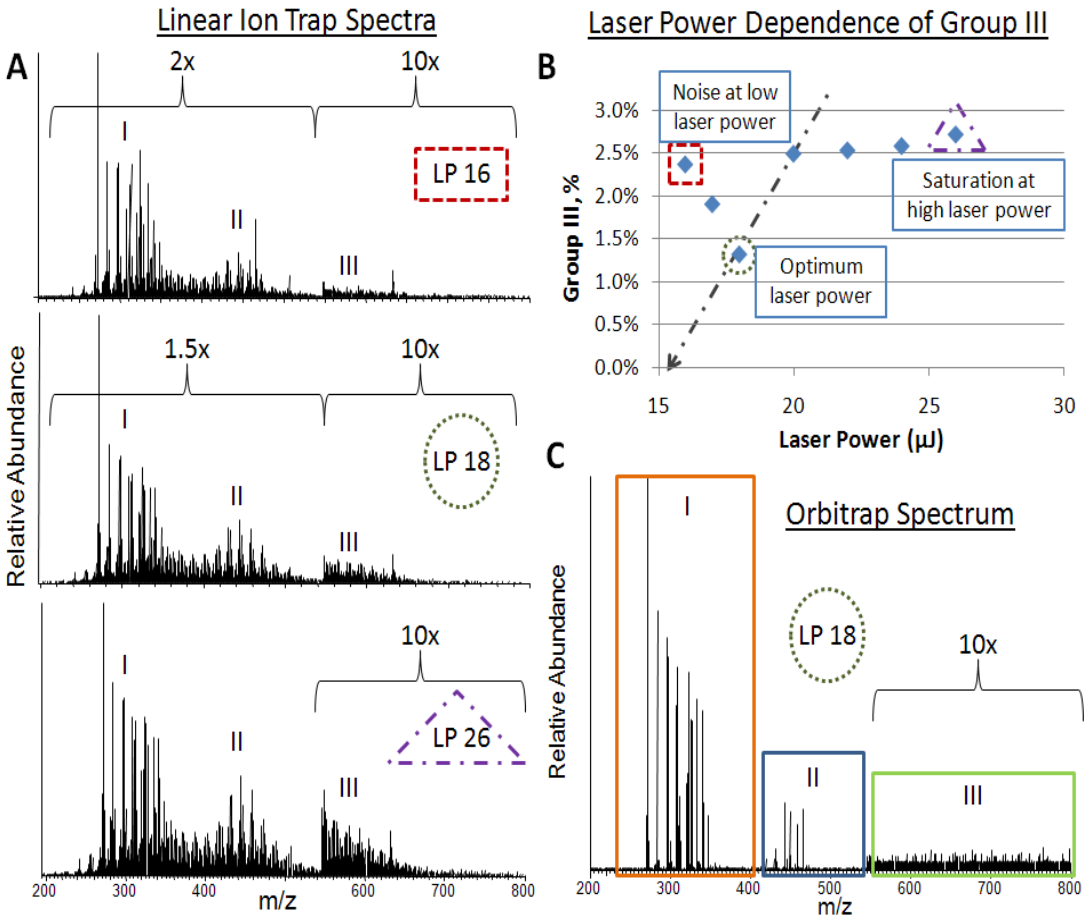


Figure 1. Laser power optimization experiment to minimize laser induced aggregation. A: Linear ion trap mass spectra at the laser power of 16, 18, and 26 μJ. Magnification is made except m/z 272. B: Relative ion signals for Group III as a function of laser power. C: Orbitrap mass spectrum at the laser power of 18 μJ.

Group I molecules are dominant, comprising ~45% of the total ion current, while Group III molecules comprise 3 %, or less, of the total ion current (Figure 1B). Relative abundances do not change much with laser power between Group I and II molecules, except slightly higher abundance of Group I at the lowest laser power of 16 μJ. The base peak of m/z 272 is unique in that its relative abundance is higher at lower laser power (note that magnification was made except at m/z 272 in Figure 1A). On the other

hand, the relative ion signals for Group III are the lowest at the laser power of 18 μJ (Figure 1B). As the laser power surpasses 18 μJ, the relative ion signal from Group III increases, most likely as a result of laser induced aggregation, and eventually becomes saturated. This suggests the aggregation reaction may be limited by sample density. At low laser power, the overall signal gets worse (total ion count of 4,290 at 16 μJ compared to 42,000 at 18 μJ) and the contribution from background noise increases resulting in higher relative abundance of Group III at low laser power (Figure 1B). Although we cannot rule out the possibility that there are some authentic bio-oil components that are not aggregation products in Group III and some aggregation products in Groups I and II, we presume Group III molecules to be aggregation products and accept Group I and II components as genuine molecules from the bio-oils. The significant signal increase of Group III components at high concentration further supports this hypothesis.

Under the above assumption, we expect that there will be no aggregation at the laser power of 15 μJ or lower (from the trend indicated by the dotted line in Figure 1B). However, ion signals get rapidly worse below 18 μJ, as discussed above, and we could not acquire high quality Orbitrap mass spectra at a laser power of 15 μJ or below. Hence, the Orbitrap mass spectrum was acquired at the optimum laser power of 18 μJ where both the background noise and aggregation products are minimized (Figure 1C). Group III is almost completely diminished in the Orbitrap spectrum at this laser power. It is suspected that most of Group III molecules are metastable, non-covalent aggregates and they might not have survived during the transport from ion trap to Orbitrap. Only

Group I and II peaks in Figure 1C were used for the subsequent data analysis and we suspect the laser induced aggregates in Group I or Group II should also be much less abundant or absent in the Orbitrap spectrum, and should be ignorable. The optimum laser power changes significantly depending on experimental conditions and sample density. Thus, the optimal laser power was determined for each sample spot before every measurement.

In addition to laser power optimization, several other preliminary experiments were performed to attain optimum experimental conditions. Several matrixes were tested for possible improvements in overall signals and/or enhancement of particular classes of molecules; however, no significant differences were observed compared to experiments that used no matrix. Slight signal improvement was noticed when colloidal graphite was used as a matrix, but there was also significant graphite background. Therefore, we obtained all the data without any matrix to simplify the data analysis and minimize the contamination from the matrix. Various solvents (methanol, 2-propanol, and acetonitrile) were tested to verify there was no reaction with the solvent, such as esterification, and no difference was noticed between the solvents used.

Chemical Composition Analysis of Bio-oil Components

External calibration provides mass accuracy of 5 ppm in Orbitrap mass spectra and internal calibration gives a slightly better mass accuracy of 3 ppm. For better mass accuracy, the experiment was performed by adding pyrene ($C_{16}H_{10}$; m/z 202.0777) to the bio-crude oil sample as an internal standard. Figure 2A and 2C show the Orbitrap spectra obtained in the m/z range 250-400 and 400-550, respectively. Pyrene was

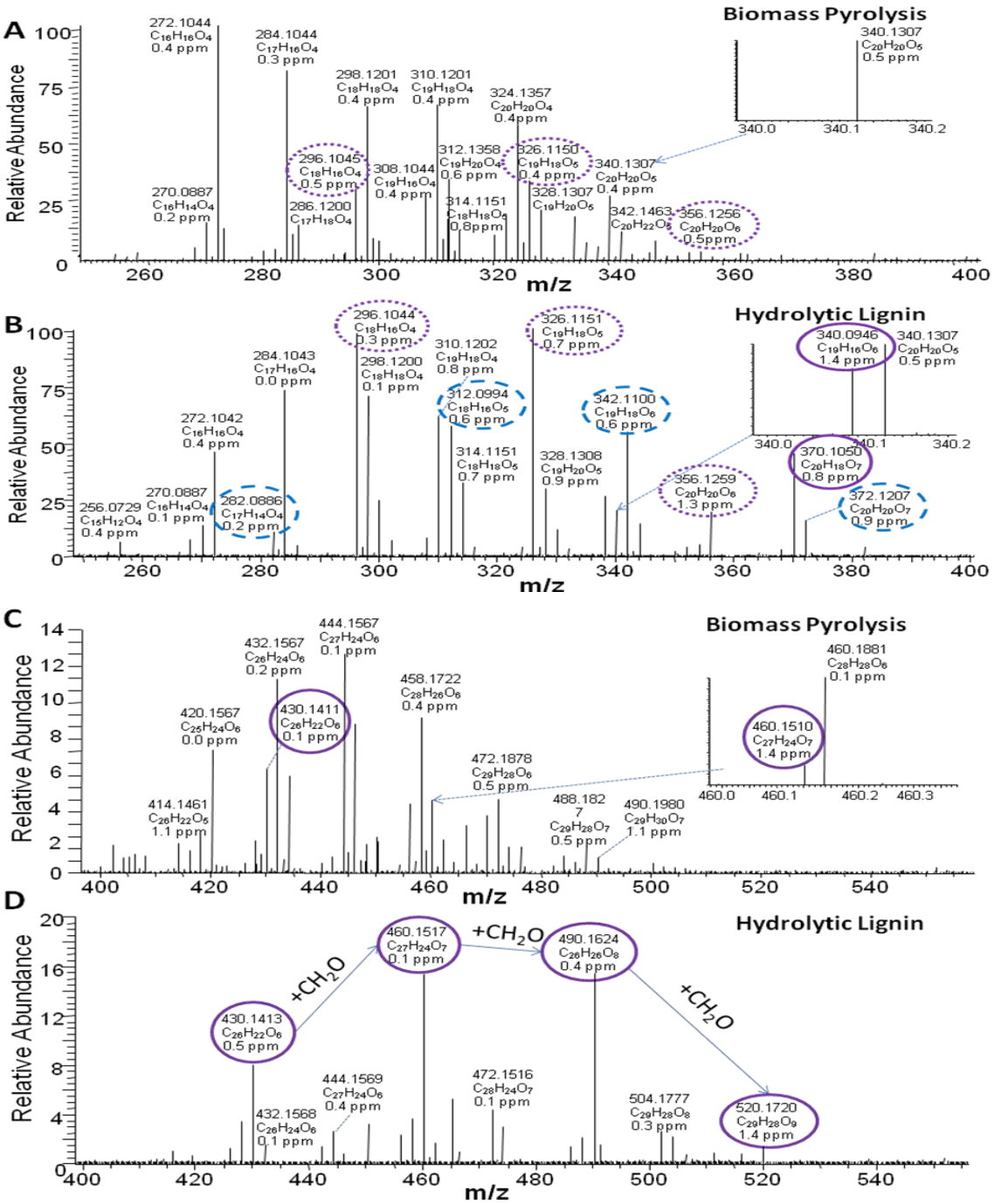


Figure 2. Orbitrap MS spectra for biomass pyrolysis sample (A, C) and hydrolytic lignin (B, D) for m/z range 250-400 (A, B) and 400-550 (C, D). Circled, dotted, and dashed lines are series of ions potentially with methoxy group addition.

detected with a mass error of 0.3 ppm with external calibration (well below m/z 250 and not shown in Figure 2A). After the internal calibration, the chemical composition analysis of the most abundant peak, m/z 272.1044, was performed with the maximum number of carbons, hydrogens, nitrogens, and oxygens of 30, 60, 10, and 10, respectively. Sulfur was ignored because the A+2 isotope for ^{34}S was not observed. Chemical composition of m/z 272.1044 was assigned as $C_{16}H_{16}O_4$ with a mass error of only 0.4 ppm and was the only chemical composition with a mass error of less than 3 ppm. All the major peaks could be assigned in the same manner and the mass errors were all less than 1 ppm. There are a few things to be noted for the major peaks shown in Figure 2A. First, they are all even mass ions, molecular radical ions with no nitrogen or proton, and oxygen is the only heteroatom. In addition, we do not see any doubly or other multiply charged ions in our data set which can be easily determined in high mass resolution spectra; i.e., doubly charged ions would produce +0.5 Da peak for the C13 isotope. Second, most of the major ions are at least 2 Da apart from each other, ignoring ^{13}C isotope, and are composed of a few series of ions with 14 Da mass differences (alkyl chain series) and 2 Da mass gaps (double bonds) between each series. The similar pattern of +14 Da and +2 Da series of ions are commonly found in petroleum oils with a much wider distribution of alkyl chains and double bonds[3, 8, 17] and also in bio-oils from biomass pyrolysis.[3,8]

Further analysis of all the ions was systematically performed by Kendrick mass defect analysis as shown in Figure 3.[18] IUPAC mass was transformed to Kendrick mass by multiplying 14.01565 (mass of CH_2) / 14. The Kendrick mass defect, the difference

between Kendrick nominal mass and Kendrick accurate mass, is consistent regardless of the size of alkyl chains for the same class (heteroatoms) and type (double bonds) molecules. After the mass normalization of alkyl chain series through Kendrick mass transformation, a series of ions with the same class and type are aligned horizontally at the same Kendrick mass defect (Figure 3). Kendrick mass defect analysis has several advantages. First, a trend of ions with the same class and type can be easily noticed. Also, low mass accuracy for high mass ions or low intensity ions can be tolerated from the trend of low mass ions and/or high abundance ions. A total of 134 compounds were assigned through direct chemical composition analysis and the Kendrick mass analysis as tabulated in Supplementary Table 1. They have heteroatom classes of O3 to O7 and DBE values of 9-17. The mass errors are all within 3 ppm except for $C_{30}H_{28}O_7$ with 4.1ppm mass error, but this is accepted based on the clear trend of other ions in that series. It should be noted that we analyzed only even mass ions in this analysis. We could not find any odd mass ions with significant intensities that are not 13C isotopes and/or contaminations.

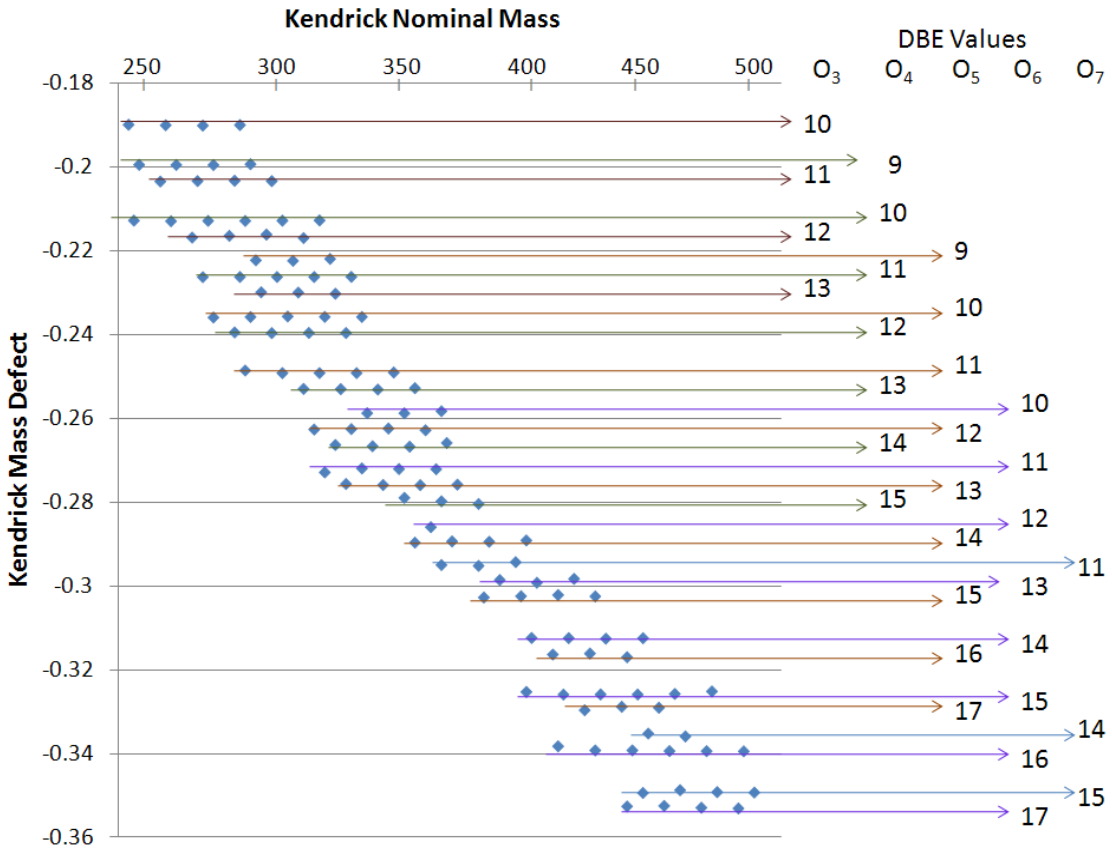


Figure 3. Kendrick mass analysis of biomass pyrolysis sample. Kendrick mass normalizes IUPAC mass with alkyl chain series. Kendrick mass defect (Kendrick nominal mass – Kendrick exact mass) allows for the confident assignment of heteroatom class and DBE.

Comparison of Fast Pyrolysis Bio-oils with Hydrolytic Lignin

To have a better understanding of the bio-oils produced by fast pyrolysis of biomass, we acquired an Orbitrap mass spectrum of hydrolytic lignin (Sigma, Catalog #: 37-107-6) at the optimal laser power obtained through the procedure described above and compared that with the fast pyrolysis of loblolly pine shown in Figure 2. According to the information provided by Sigma, this sample is a polymeric autohydrolysis lignin isolated from a commercial hydrolysis pilot plant using predominantly sugar cane

bagasse as a raw material. It is in solid form and has average molecular weight of 1,700 (number averaged molecular weight) or 19,300 (weight averaged molecular weight) according to the data sheet. As shown in Figure 2B and 2D, however, only small molecules in m/z range 250-550 could be observed (no peak above m/z 550 which is not shown). It seems we detect only low molecular weight components of hydrolytic lignin under the given experimental conditions.

Hydrolytic lignin has two groups of peaks similar to Group I and II in biomass pyrolysis (Figures 2B and 2D). Spectral similarity could be noticed with biomass pyrolysis, especially in Group I (Figures 2A and 2B). Many of the major peaks in Group I have exactly the same accurate masses (within mass error of 1ppm) and corresponding chemical compositions between biomass pyrolysis and hydrolytic lignin. The Group II, however, shows quite different mass spectral patterns (Figures 2C and 2D). Hydrolytic lignin is dominated by a series of peaks with mass difference of ~30.0104 Da; m/z 430.1413, 460.1517, 490.1624, and 520.1720. This mass difference corresponds to the chemical composition of CH_2O suggesting that they might have originated from the same chemical compound that was modified with a different number of methoxy groups. Biomass pyrolysis also has m/z 460.1510, a possible methoxy group addition to m/z 430.1411, but in very low abundance.

It is not as outstanding as in Group II, but Group I also shows some evidence of molecules with high methoxy group contents in hydrolytic lignin. For example, m/z 340.0946 ($C_{19}H_{16}O_6$) and m/z 370.1050 ($C_{20}H_{18}O_7$) are only in hydrolytic lignin (Figure 2B; solid circled). Two other methoxy group series could be clearly noted in Figure 2B

(dotted and dashed circles, respectively): m/z 296.1044 ($C_{18}H_{16}O_4$), m/z 326.1151 ($C_{19}H_{18}O_5$) and m/z 356.1259 ($C_{20}H_{20}O_6$); m/z 282.0886 ($C_{17}H_{14}O_4$), m/z 312.0994 ($C_{18}H_{16}O_5$), m/z 342.1100 ($C_{19}H_{18}O_6$), and m/z 372.1207 ($C_{20}H_{20}O_7$). It should be noted that both the series of ions are also observed in biomass pyrolysis, but in lower abundance (dotted circles in Figure 2A). Careful inspection of Supplementary Table 1 revealed that most of the O5 and O6 compounds in Group I can be explained as CH_2O addition to O4 and O5 compounds, respectively, and all O7 compounds in Group II can be explained as CH_2O addition to O6 compounds. It may indicate that, contrary to the apparent complexity of bio-oils, they have originated from a few common structures that are extended with different levels of alkylation, double bond formations and methoxy group additions.

The pine lignin is known to have a methoxy group attached to an aromatic ring and bagasse has one or two methoxy groups[22], which is also consistent with our observation that bagasse acidic hydrolysis has higher abundance in methoxy groups than pine pyrolysis, suggesting the structural influence of original biomass materials. A systematic study on various biomass materials and chemical processes may be needed for the better understanding how they affect the final products.

Understanding Petroleomic Data Set

Various graphical representations have been developed to illustrate thousands of molecular components in petroleum oils obtained from petroleomics approaches.[23] There is a fundamental assumption in all this data exploration that we ignore the

difference in ionization yields between the compounds. This assumption has a clear limitation because different functional groups may result in significant differences in ionization yields. However, no method has been reported so far that can quantify hundreds and/or thousands of compounds in a complex mixture of petroleum oils or bio-oils, most of which exact chemical structures are still unknown. We adopt the assumption of difference in ionization yield as a zeroth-order approximation and utilize some graphical tools developed in petroleomics to illustrate and understand hundreds of molecular components in bio-oils. Hence, the data presented here will provide only qualitative and/or semi-quantitative information.

It should also be noted that there are certain classes of molecules missing in the current study. First, as mentioned above, volatile compounds, with vapor pressure higher than ~80 mtorr (our MALDI source pressure), will evaporate rapidly from the sample plate before mass spectral data acquisition. Second, we may preferentially ionize those compounds that have high ionization yields under LDI condition. Hage and coworkers demonstrated multi-photon ionization could be efficiently used to ionize pyrolyzed lignin compounds.[12] Particularly, they obtained a spectral pattern comparable to ours in pyrolyzed wood oak sample for m/z range 250-400 (m/z 272, 284, 302, 314, 344, 332 and also ±2Da shifted ions around these ions) using a laser wavelength (354.6 nm) similar to ours (337.7 nm), which suggests the ions in our data set might also be produced via multi-photon ionization by favoring aromatic compounds and discriminating against other compounds, such as cellulose pyrolysis products. The similarity of our current data on biomass pyrolysis with lignin hydrolysis is partially

owing to the fact that the pyrolytic products of pine cellulose might not have been readily ionized and detected with the current method.

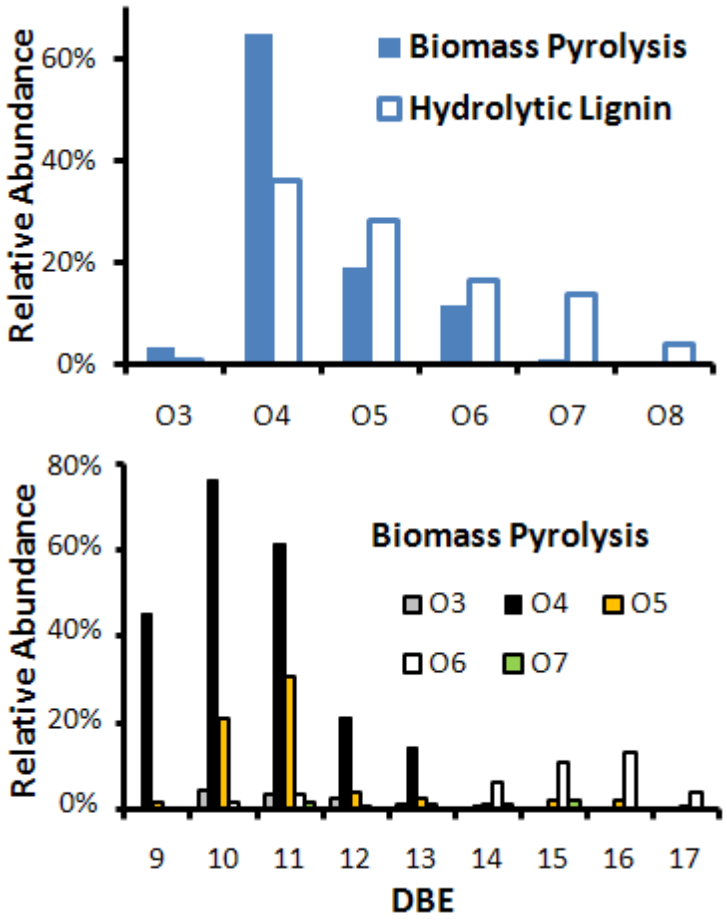


Figure 4. A: Heteroatom class analysis for biomass pyrolysis and hydrolytic lignin. B: DBE distribution of biomass pyrolysis sample for each heteroatom class.

Figure 4A shows the abundance of each heteroatom class with comparisons between the biomass pyrolysis and the hydrolytic lignin. In both samples, O4 compounds are dominant, followed by O5 and O6 compounds; however, O4 compounds are only 36% in hydrolytic lignin while they are 64% in biomass pyrolysis. It has been

known that minimum oxygen content in biodiesel is important to maintain long term storage.[24] Higher oxygen content in hydrolytic lignin seems to come from a higher number of methoxy groups as discussed above. Figure 4B shows the plot of double-bond equivalents (DBE) distribution for each heteroatom class for the fast pyrolysis of biomass. Two DBE distributions, centered at DBE of 10-11 and 15-16, respectively, are distinctively noticed. The first distribution is dominated by O4 and O5 compounds, and the second comes mostly from O6 compounds. The high DBE values of 9-17 suggest they most likely contain at least one aromatic group, but possibly two or three.

The molecular components we identified in the bio-oil of fast-pyrolysis biomass seem to be de-polymerized products of lignins, particularly dimeric and trimeric units of lignin for the 1st and 2nd DBE distributions in Figure 4B. First of all, high similarity in chemical compositions between fast pyrolysis biomass and hydrolytic lignin strongly supports this hypothesis. De-polymerized lignin has been suggested as a common pyrolysis product of wood materials.[25,26] Also, the particular fraction used for the current study is highly water-insoluble and expected to contain a high concentration of lignin.[27] It is noteworthy that lignin monomers are missing in our data set because they are volatile with m/z below 250 while other studies using different types of ionization methods could detect lignin monomers such as in pyrolysis GC-MS[28], pyrolysis FI-MS analysis of lignins[10, 15], ESI-MS[3], and GC-MS[4]. Some other supportive evidence can be found in the similarity of the weight % ratio of C/H/O between lignin and our data. The C:H:O ratios of 69:5:26 in hypothetic lignin and 72:5:23 in hydrolytic lignin are similar to that of our data set, 76: 6: 18. Here, the hypothetical lignin is calculated from a

theoretical model of lignin nanomer proposed by Bayerbach and Meier.[9] Lower oxygen content in fast pyrolysis biomass, compared to hypothetical lignin or hydrolytic lignin, might originate from the loss of some oxygen during the pyrolysis process, most likely through the loss of some methoxy groups. It should be noted that molecular characteristics is often lost in pyrolysis at very high temperatures; for example, FT-IR spectra of lignin and cellulose pyrolysis at 600 °C or higher are very similar[29]. However, we expect quite a portion of molecular characteristics would be still remained for the pyrolysis at 500 °C or below, such as the bio-oil sample used in the current study.

Figure 5 shows the contour map of O4 and O6 compounds, the most dominant heteroatom species in dimeric and trimeric units of lignin, displayed for the number of carbon and DBE as x- and y-coordinates. The area of each circle in this contour diagram represents the ion signals for each corresponding molecule. There are two distinct contours: one centered at #C of ~18 and DBE of ~10 with O4 compounds dominating, and the other centered at #C of ~27 and DBE of ~15 with O6 compounds dominating. The average difference between the two is 9 for the number of carbons, 2 for the number of oxygens, and 5 for DBE, implying the average chemical composition of the monomer unit might be $C_9H_{10}O_2$. This strongly suggests that the two contours correspond to dimers and trimers of lignin because the average chemical compositions of the two correspond to $C_{18}H_{18}O_4$ and $C_{27}H_{26}O_6$, exactly two and three times the monomer chemical compositions except one or two H_2 loss, respectively, for the conjugation between the monomer units. Average chemical composition of the monomer unit, $C_9H_{10}O_2$, happens to be the same with coumaryl alcohol (HO-Ph-CH=CH-CH_2OH), one

of three possible lignin monomers, further supporting the lignin hypothesis; however, the structure is most likely not exactly the same considering coumaryl alcohol is not a major monomer in pine lignin but it should be something produced through pyrolytic depolymerization. It should be noted that there is a difference between 'lignin monomer' as a building block of polymeric lignin in plant material and 'monomeric unit' in the final products of biomass pyrolysis. The latter is originated from the former, so it should have structural similarity, but pyrolytic process should have made it structurally diversified.

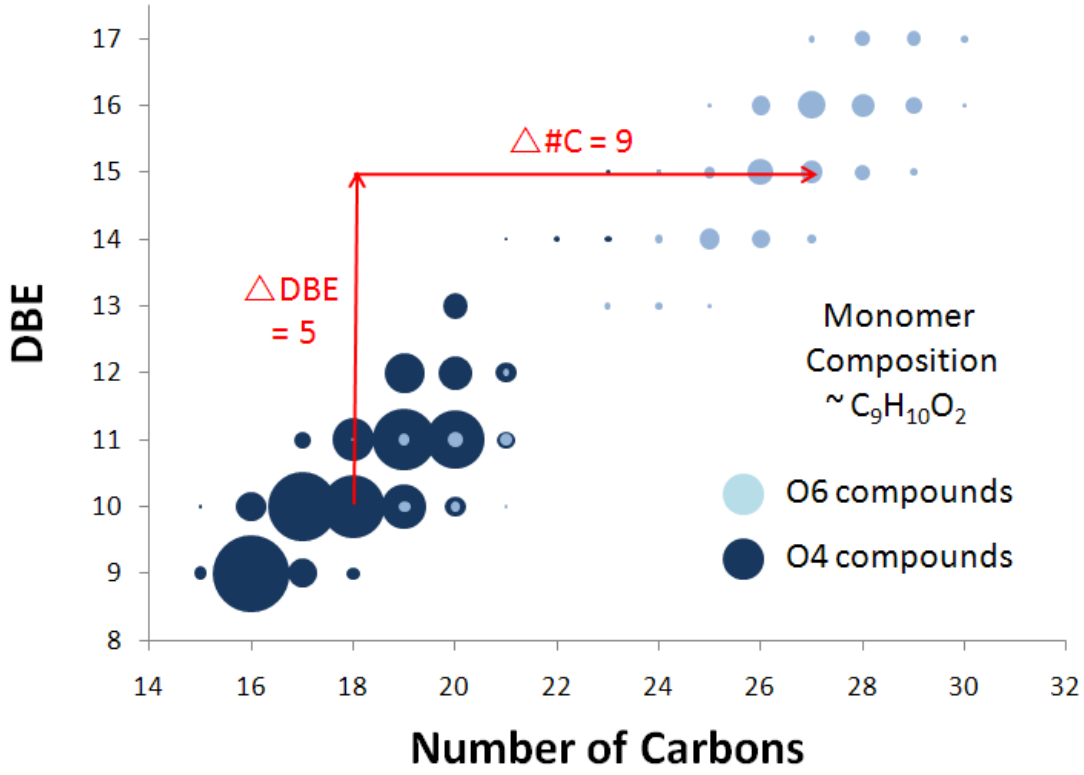


Figure 5. Contour map of number of carbons vs. DBE for O4 and O6 compounds of biomass pyrolysis sample. The dark blue represents O4 heteroatom class and the light blue represents O6 heteroatom class with the area of each circle representing intensity of corresponding molecule.

Structural Insights from MS/MS Analysis

To obtain structural details, we performed MS/MS analysis of some of the chemical compounds in fast pyrolysis biomass. As discussed above, the bio-oils are much less complex than petroleum oils and many of the major compounds have almost no interfering peaks within ±2 Da, other than 13C isotopes. This is different from petroleum oils, where at least a few peaks can be found with significant intensities within every Da. This provides us an exclusive opportunity to perform MS/MS of some precursor ions with almost no interference from nearby peaks, which is impossible in petroleum oil analysis. It should be noted that some peaks do have interference at the same nominal mass and cannot be clearly isolated for MS/MS. For example, m/z 326.1515 ($C_{20}H_{22}O_4$) has an interference of m/z 326.1150 ($C_{19}H_{18}O_5$) with 22 % relative intensity (Supplementary Table 1). Most other major precursor ions, however, can be isolated with no or a negligible amount of interference.

Figure 6 shows MS/MS spectra of m/z 272, 284, 296, and 308 which represent O4 compounds with DBE of 9, 10, 11, and 12, respectively. MS/MS spectra of m/z 284, 296, and 308 show very similar spectral patterns: methyl loss is most dominant, and CH_4O and C_2H_3O losses are consistently found with significant intensities. On the other hand, MS/MS of m/z 272 is clearly distinguished from the others by having much less methyl loss and many other fragmentations instead. Fragmentations corresponding to the loss of CH_3O and CH_5O, instead of CH_4O, are quite intriguing and the losses of OH and $C_2H_3O_2$ are also very unique. Typically, water loss is observed instead of OH for a hydroxyl group attached to an alkyl chain, indicating OH loss in MS/MS of m/z 272

might be from a phenolic hydroxyl group. The loss of CH_3O and CH_4O may have originated from the direct cleavage or rearrangement fragmentation of a methoxy group, and is consistent with the high content of methoxy groups known to be present in lignin and/or bio-oils suggested by IR and NMR study.[3] The precursor ion of m/z 272 is very abundant in the MS/MS spectrum, five times or higher than any fragment, while other precursors are less abundant than the base fragment of methyl loss, suggesting unusual stability of $C_{16}H_{16}O_4^{+\bullet}$ molecular radical ion.

The uniqueness of m/z 272 can also be noted elsewhere. First, as shown in Figure 5, carbon distribution is very narrow for the DBE of 9, centered at 16 carbons ($C_{16}H_{16}O_4$) constituting 80% of all O4 compounds with DBE 9, while other DBE groups have much wider carbon distributions with any carbon number being less than 40% within each DBE group. Second, m/z 272 is very distinctively observed in others' studies. It was observed as the most dominant peak above m/z 250, regardless of biomass or pretreatment by Evans and coworkers[19] and regardless of electron energy by Xu and coworkers[8]. The peak at m/z 272 is even observed in pyrolysis FI-MS of forest humus samples by Hempfling and Schulten and suggested to have come from the pyrolysis of lignin.[15] The anomalous MS/MS spectrum, combined with the other unusual behavior of m/z 272, strongly suggests that m/z 272 must have a very different chemical structure compared to the others. Considering the possibility that m/z 272 could be a mixture of several structural isomers with at least the major isomer having a very unique structure.

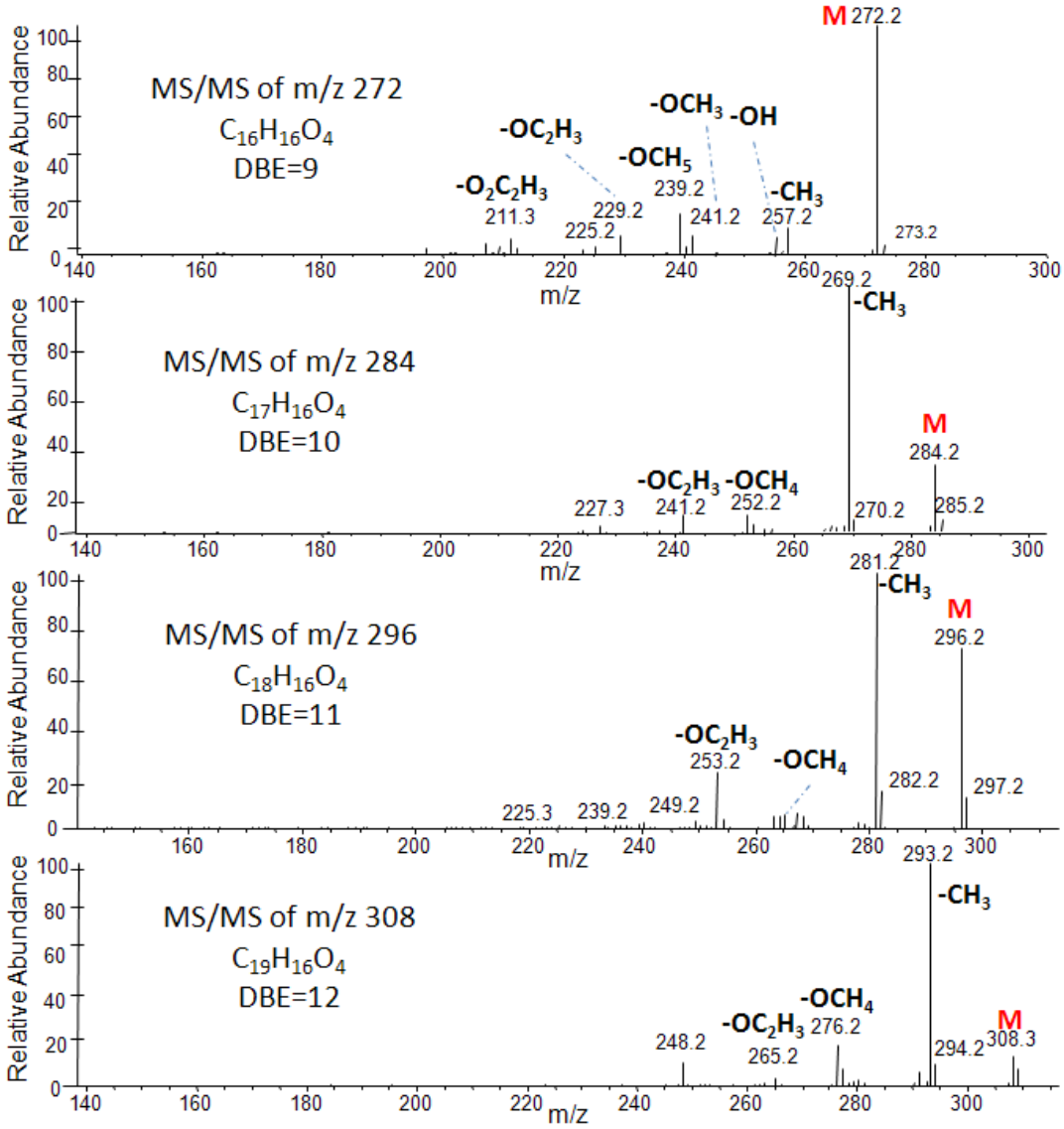


Figure 6. MS/MS spectra of m/z 272, 284, 296, and 308; representative O4 compounds with DBE of 9, 10, 11, and 12, respectively.

It is noteworthy that the chemical composition of m/z 272 is also assigned as $C_{16}H_{16}O_4$ in the pyrolysis of lignin by Hempfling and Schulten[15] and by Evans and coworkers[9, 19]. Hempfling and Schulten determined the exact chemical composition from high resolution mass spectra acquired with a double focusing mass spectrometer

while Evans' group simply speculated. Evans' group suggested it has an alkylarylether linkage between the two aromatic monomer units, but Hempfling and Schulten claimed that alkylaryehther would be thermally labile and would not survive in the harsh pyrolysis process. Instead, they suggested two alternative structures of phenyl coumaran and a stilben derivative. Nonetheless, both the groups do not have any data supporting their suggested structures. Bayerbach and Meier suggested a diphenylether type structure according to their FTIR and NMR data[9]; however, their FTIR and NMR measurements were of a mixture of biomass pyrolysis compounds, and not necessarily representative of a particular compound, such as m/z 272. A detailed study on MS/MS of various standards and/or model compounds will be needed to get further insights on their structures.

Conclusion

The use of Orbitrap high resolution mass spectrometry with laser desorption ionization allowed for the identification of 136 chemical compositions of non-volatile compounds in a bio-oil sample produced by fast pyrolysis of loblolly pine. Comparison with low molecular weight components in hydrolytic lignin and petroleomic analysis suggests these compounds are dimer and trimer units of lignin de-polymerization products. Our petroleomic analysis predicts average chemical compositions of monomeric unit as $C_9H_{10}O_2$. Partial structural information has been obtained from MS/MS of a few molecular components.

The choice of laser desorption ionization in the current study has limited us to non-volatile aromatic compounds. For example, we failed to achieve meaningful signals in our preliminary effort for cellulose pyrolysis using the current approach. It is attributed to the inability of the non-aromatic cellulose pyrolysis products to absorb the laser power for their desorption and ionization. We have a plan to adopt other ionization methods, specifically ESI and APCI, to analyze the pyrolytic products of cellulose and hemicellulose. However, the current approach adopted here turned out to be well suited for the first application of petroleomic analysis of fast pyrolysis bio-oils. According to our recent preliminary experiment using APPI FTICR, APPI also produces mostly lignin dimeric and trimeric units, but with much more complexity due to the overlap between protonated molecules and molecular radical ions. Further investigation with FTICR using various ionizations including APPI, ESI, and APCI will provide us a better understanding of fast pyrolysis bio-oils. A systematic MS/MS study will also be performed for major compounds and various standard compounds to attain detailed structural insights. We are currently working on a project to study bio-oil aging in which not only the current tool but also other traditional analytical techniques, such as GC-MS and gel permeation chromatography, will be used to achieve comprehensive understanding of biomass pyrolysis products.

Acknowledgement

This work was supported by grants from Iowa State University and Ames Laboratory-USDOE, and partially supported by ConocoPhillips. The authors thank Robert C. Brown, Center for Sustainable Environmental Technology at Iowa State University, for bio-oil samples and valuable discussions and Christopher J. Thompson at Bruker for LDI-FT ICR analysis of bio-oils. EAS acknowledges GAANN fellowship from U.S. Department of Education.

References

1. T. L. Marker, et al., ed. P. b. U. U.S. Department of Energy Final Report, Inc.,, 2005.
2. E.-b. Hassan, P. Steele and L. Ingram, *Applied Biochemistry and Biotechnology*, 2009, **154**, 3-13.
3. G. r. Gellerstedt, J. Li, I. Eide, M. Kleinert and T. Barth, *Energy & Fuels*, 2008, **22**, 4240-4244.
4. M. Kleinert and T. Barth, *Energy & Fuels*, 2008, **22**, 1371-1379.
5. A. Oasmaa and S. Czernik, *Energy & Fuels*, 1999, **13**, 914-921.
6. G. Jiang, D. J. Nowakowski and A. V. Bridgwater, *Energy & Fuels*, 2010, null-null.
7. S. Kim, R. P. Rodgers, G. T. Blakney, C. L. Hendrickson and A. G. Marshall, *Journal of the American Society for Mass Spectrometry*, 2009, **20**, 263-268.
8. F. Xu, Y. Xu, H. Yin, X. Zhu and Q. Guo, *Energy & Fuels*, 2009, **23**, 1775-1777.
9. R. Bayerbach, V. D. Nguyen, U. Schurr and D. Meier, *Journal of Analytical and Applied Pyrolysis*, 2006, **77**, 95-101.
10. R. J. Evans, T. A. Milne and M. N. Soltys, *Journal of Analytical and Applied Pyrolysis*, 1986, **9**, 207-236.

11. A. L. Brown, D. C. Dayton, M. R. Nimlos and J. W. Daily, *Chemosphere*, 2001, **42**, 663-669.

12. E. R. E. van der Hage, J. J. Boon, R. J. J. M. Steenvoorden and T. L. Weeding, *Analytical Chemistry*, 1994, **66**, 543-550.

13. A. R. Hortal, P. Hurtado, B. Martinez-Haya and O. C. Mullins, *Energy & Fuels*, 2007, **21**, 2863-2868.

14. A. E. Pomerantz, M. R. Hammond, A. L. Morrow, O. C. Mullins and R. N. Zare, *Energy & Fuels*, 2009, **23**, 1162-1168.

15. R. Hempfling and H. R. Schulten, *Organic Geochemistry*, 1990, **15**, 131-145.

16. T. M. Schaub, C. L. Hendrickson, S. Horning, J. P. Quinn, M. W. Senko and A. G. Marshall, *Analytical Chemistry*, 2008, **80**, 3985-3990.

17. A. G. Marshall and R. P. Rodgers, *Accounts of Chemical Research*, 2004, **37**, 53-59.

18. C. A. Hughey, C. L. Hendrickson, R. P. Rodgers, A. G. Marshall and K. Qian, *Analytical Chemistry*, 2001, **73**, 4676-4681.

19. R. J. Evans and T. A. Milne, *Energy & Fuels*, 1987, **1**, 123-137.

20. A. Sherwood-Pollard, M.S. M.S. Thesis, Iowa State University, 2009.

21. K. Strupat, V. Kovtoun, H. Bui, R. Viner, G. Stafford and S. Horning, *Journal of the American Society for Mass Spectrometry*, 2009, **20**, 1451-1463.

22. G. De Stevens and F. F. Nord, *PNAS*, 1953, **39**, 80-84.

23. A. Marshall and R. Rodgers, *PNAS*, 2008, **105**, 18090-18095.

24. J. P. Diebold, *A review of the chemical and physical mechanisms of the storage stability of fast pyrolysis bio-oils*, Report NREL/SR-570-27613; TRN: AH200011%%11, 1999.

25. D. Mohan, C. U. Pittman and P. H. Steele, *Energy & Fuels*, 2006, **20**, 848-889.

26. A. V. Bridgwater and G. V. C. Peacocke, *Renewable and Sustainable Energy Reviews*, 2000, **4**, 1-73.

27. A. G. Gayubo, B. Valle, A. T. Aguayo, M. Olazar and J. Bilbao, *Journal of Chemical Technology & Biotechnology*, 2010, **85**, 132-144.

28. B. Scholze and D. Meier, *Journal of Analytical and Applied Pyrolysis*, 2001, **60**, 41-54.

29. N. L. Wang, M.J.D, *Materials Chemistry and Physics*, 1990, **26**, 67-80.

Supplementary Table 1. Chemical compositions of the bio-oil produced from fast pyrolysis of biomass.

O_3 Compounds				
Formula	$Mass_{exp}$	$Mass_{cal}$	Δm, ppm	Intensity
DBE = 9				
C16 H16 O3	256.1095	256.1094	0.4	5405
DBE=10				
C16H14O3	254.0939	254.0937	0.6	8936
C17H16O3	268.1094	268.1094	0.2	20661
C18H18O3	282.1250	282.1250	-0.2	12277
C19H20O3	296.1408	296.1407	0.3	2522
DBE=11				
C17H14O3	266.0937	266.0937	0.0	3024
C18H16O3	280.1095	280.1094	0.3	15275
C19H18O3	294.1252	294.1250	0.6	12733
C20H20O3	308.1408	308.1407	0.3	3165
DBE=12				
C18H14O3	278.0937	278.0937	0.0	6368
C19H16O3	292.1098	292.1094	1.4	9876
C20H18O3	306.1258	306.1250	2.3	6051
C21H20O3	320.1406	320.1407	-0.4	2209
DBE=13				
C20H16O3	304.1097	304.1094	1.1	2559
C21H18O3	318.1253	318.1250	0.8	4808
C22H20O3	332.1407	332.1407	0.0	1198

O_4Compounds				
Formula	$Mass_{exp}$	$Mass_{cal}$	Δm, ppm	Intensity
DBE=9				
C15H14O4	258.0888	258.0887	0.5	11779
C16H16O4	272.1044	272.1043	0.4	373953
C17H18O4	286.1200	286.1200	0.3	56675
C18H20O4	300.1359	300.1356	0.9	10015
DBE=10				
C15H12O4	256.0732	256.0730	0.8	1113
C16H14O4	270.0887	270.0887	0.2	59690
C17H16O4	284.1044	284.1043	0.3	302598
C18H18O4	298.1201	298.1200	0.4	245708
C19H20O4	312.1358	312.1356	0.6	129555
C20H22O4	326.1515	326.1513	0.7	27242
DBE=11				
C17H14O4	282.0888	282.0887	0.3	17881
C18H16O4	296.1045	296.1043	0.5	113224
C19H18O4	310.1201	310.1200	0.4	248067
C20H20O4	324.1357	324.1356	0.4	215889
C21H22O4	338.1514	338.1513	0.5	18466
DBE=12				
C18H14O4	294.0890	294.0887	1.0	9467
C19H16O4	308.1044	308.1043	0.4	102155
C20H18O4	322.1201	322.1200	0.5	73051
C21H20O4	336.1358	336.1356	0.5	27231
DBE=13				
C20H16O4	320.1045	320.1043	0.6	39734
C21H18O4	334.1201	334.1200	0.5	68845
C22H20O4	348.1357	348.1356	0.1	30110
C23H22O4	362.1516	362.1513	1.0	2384
DBE=14				
C21H16O4	332.1045	332.1043	0.7	1074
C22H18O4	346.1199	346.1200	-0.3	2988
C23H20O4	360.1354	360.1356	-0.5	2836
C24H22O4	374.1519	374.1513	1.8	1560
DBE=15				
C23H18O4	358.1209	358.1200	2.8	1064
C24H20O4	372.1358	372.1356	0.4	1146
C25H22O4	386.1508	386.1513	-1.2	1866

O_5 Compounds				
Formula	$Mass_{exp}$	$Mass_{cal}$	Δm, ppm	Intensity
DBE=9				
C17H18O5	302.1151	302.1149	0.9	10059
C18H20O5	316.1307	316.1305	0.4	3005
C19H22O5	330.1468	330.1462	1.8	3358
DBE=10				
C16H14O5	286.0836	286.0836	0.0	5820
C17H16O5	300.0994	300.0992	0.4	30547
C18H18O5	314.1151	314.1149	0.8	47751
C19H20O5	328.1307	328.1305	0.5	80349
C20H22O5	342.1463	342.1462	0.4	46097
DBE=11				
C17H14O5	298.0843	298.0836	2.5	1669
C18H16O5	312.0994	312.0992	0.4	65408
C19H1805	326.1150	326.1149	0.4	126191
C20H20O5	340.1307	340.1305	0.4	102688
C21H22O5	354.1464	354.1462	0.8	11881
DBE=12				
C19H16O5	324.0993	324.0992	0.1	5077
C20H18O5	338.1150	338.1149	0.5	21772
C21H20O5	352.1308	352.1305	0.8	7232
C22H2205	366.1460	366.1462	-0.4	2669
DBE=13				
C20H16O5	336.0997	336.0992	1.3	2225
C21H18O5	350.1151	350.1149	0.6	7779
C22H20O5	364.1307	364.1305	0.5	8911
C23H22O5	378.1465	378.1462	0.8	3088
DBE=14				
C22H18O5	362.1147	362.1149	-0.5	1390
C23H20O5	376.1306	376.1305	0.3	2574
C24H22O5	390.1462	390.1462	0.1	4277
C25H24O5	404.1622	404.1618	0.9	3407
DBE=15				
C24H20O5	388.1307	388.1305	0.4	2617
C25H22O5	402.1466	402.1462	1.1	6304
C26H24O5	416.1625	416.1618	1.7	5176
C27H26O5	430.1779	430.1775	0.9	2759
DBE=16				
C26H22O5	414.1461	414.1462	-0.2	6729
C27H24O5	428.1620	428.1618	0.4	7272
C28H26O5	442.1767	442.1775	-1.6	3611
DBE=17				
C27H22O5	426.1462	426.1462	0.0	2060
C28H24O5	440.1627	440.1618	1.9	2129
C29H26O5	454.1781	454.1775	1.3	1822

O_6 Compounds				
Formula	$Mass_{exp}$	$Mass_{cal}$	Δm, ppm	Intensity
DBE = 9				
C18 H20 O6	332.1253	332.1254	-0.4	4477.7
DBE=10				
C18H18O6	330.1102	330.1098	1.2	982.5
C19H20O6	344.1255	344.1254	0.3	9231
C20H22O6	358.1412	358.1411	0.2	6020
C21H2406	372.1573	372.1567	1.5	1305
DBE=11				
C18H16O6	328.0935	328.0941	-1.9	1051
C19H18O6	342.1101	342.1098	0.9	8803
C20H20O6	356.1256	356.1254	0.5	13806
C21H22O6	370.1412	370.1411	0.3	8092
DBE=12				
C21H20O6	368.1251	368.1254	-0.9	3333
DBE=13				
C23H22O6	394.1416	394.1411	1.2	2586
C24H24O6	408.1565	408.1567	-0.5	3808
C25H26O6	422.1731	422.1724	1.7	1358
DBE=14				
C24H22O6	406.1411	406.1411	0.0	4138
C25H24O6	420.1567	420.1567	0.0	27919
C26H26O6	434.1722	434.1724	-0.3	21913
C27H28O6	448.1880	448.1880	-0.1	6536
DBE=15				
C24H20O6	404.1260	404.1254	1.3	1255
C25H22O6	418.1410	418.1411	-0.3	8764
C26H24O6	432.1567	432.1567	-0.2	44117
C27H26O6	446.1723	446.1724	-0.2	33956
C28H28O6	460.1881	460.1880	0.1	16454
C29H30O6	474.2043	474.2037	1.4	3711
DBE=16				
C25H20O6	416.1264	416.1254	2.3	1489
C26H22O6	430.1411	430.1411	-0.1	23737
C27H24O6	444.1567	444.1567	-0.1	49826
C28H26O6	458.1722	458.1724	-0.4	35325
C29H28O6	472.1878	472.1880	-0.5	16667
C30H30O6	486.2034	486.2037	-0.5	1738
DBE=17				
C27H22O6	442.1411	442.1411	-0.1	3675.5
C28H24O6	456.1568	456.1567	0.2	15753
C29H25O6	470.1721	470.1724	-0.7	13074
C30H28O6	484.1875	484.1880	-1.0	3938.8
DBE = 18				
C29H24O6	468.1569	468.1567	0.3	1960.9
C30H26O6	482.1716	482.1724	-1.6	1505.7

O_7 Compounds				
Formula	$Mass_{exp}$	$Mass_{cal}$	Δm, ppm	Intensity
DBE = 10				
C21H24O7	388.1523	388.1517	1.7	1499.9
DBE=11				
C20H20O7	372.1205	372.1204	0.5	2796
C21H22O7	386.1360	386.1360	0.0	10298
C22H2407	400.1525	400.1517	2.2	1296
DBE=14				
C26H26O7	450.1674	450.1673	0.3	8230
C27H28O7	464.1824	464.1830	-1.2	2462
DBE=15				
C26H24O7	448.1510	448.1517	-1.5	2673.5
C27H26O7	462.1673	462.1673	0.0	7393
C28H28O7	476.1825	476.1830	-1.0	5811
C29H30O7	490.1980	490.1986	-1.1	3365
C30H32O7	504.2147	504.2143	0.8	1071
DBE = 16				
C26H22O7	446.1354	446.1360	-1.4	1307.5
C27H24O7	460.1510	460.1517	-1.4	3029.6
C28H26O7	474.1677	474.1673	0.8	5844.3
C29H28O7	488.1827	488.1830	-0.5	7092.3
C30H30O7	502.1991	502.1986	1.0	1381.9
DBE = 17				
C29H26O7	486.1669	486.1673	-0.8	2395.1
C30H28O7	500.1809	500.1830	-4.1	2172.2

* Only even mass ions with intensities higher than 1000 counts were analyzed.
* No odd mass ions with significant intensities has been found which is not 13C isotope or contaminations.
* Yellow colored O5 compounds could be methoxy group addition to gray colored O4 compounds.

CHAPTER III

BIO-OIL ANALYSIS USING NEGATIVE ELECTROSPRAY IONIZATION: COMPARATIVE STUDY OF HIGH RESOLUTION MASS SPECTROMETERS AND PHENOLC VS. SUGARIC COMPONENTS

A paper published in *Energy and Fuels*

Erica A. Smith, Soojin Park, Adam T. Klein, and Young Jin Lee*
Department of Chemistry, Iowa State University and
Ames Laboratory, US-DOE, Ames, Iowa 50011

Abstract

We have previously demonstrated that a petroleomic analysis could be performed for bio-oils and revealed the complex nature of bio-oils for the non-volatile phenolic compounds (Energy & Fuels, 24, 5190-5198, 2010). As a subsequent study, we have adapted electrospray ionization in negative ion mode to characterize a wide variety of bio-oil compounds. A comparative study of three common high resolution mass spectrometers was performed to validate the methodology and to investigate the differences in mass discrimination and resolution. The mass spectrum, in negative ESI, is dominated by low mass compounds with m/z of 100-250 with some compounds being analyzable by GC-MS. We could characterize over eight hundred chemical compositions with only about forty of them being previously known in GC-MS. This unveiled a much more complex nature of bio-oils than typically shown by GC-MS. The pyrolysis products of cellulose and hemicellulose, particularly poly-hydroxy cyclic hydrocarbons or 'sugaric' compounds such as levoglucosan, could be effectively characterized with this approach. Phenolic compounds from lignin and holocellulose

pyrolysis products could be clearly distinguished in a contour map of DBE vs. the number of carbons from these sugaric compounds.

Introduction

Fast pyrolysis of biomass is a promising method to replace fossil fuels for our transportation needs and other petroleum products. Current chemical analysis for fast pyrolysis bio-oils is mostly conducted by GC-MS, FTIR, and NMR; but these methods are insufficient for a complete understanding of the molecular complexity in bio-oils.[1-3] In the last decade, modern mass spectrometry techniques have been successfully utilized for the analysis of petroleum crude oils. High-resolution mass spectrometry (HRMS) has proven to be a powerful tool for the detailed molecular characterization of crude oils with direct chemical composition analysis of over thousands of molecular components.[4] We have recently adapted a petroleomics approach for bio-oil analysis and have demonstrated the complex nature of lignin pyrolysis products at a molecular level by identifying over one hundred chemical compositions.[5] The methodology we utilized in the previous study (laser desorption ionization and orbitrap mass spectrometer) has a few inherent analytical limitations. The present work is a subsequent study in an effort to overcome some of the previous limitations and to obtain a better understanding of bio-oils.

The first aspect of the current study is comparing three commonly employed high-resolution mass spectrometers for the analysis of complex bio-oil samples. HRMS analysis of crude oils has been almost exclusively performed with Fourier transform ion

cyclotron resonance mass spectrometers (FTICR) because of its superior mass resolving power needed for the enormous complexity of crude oils. However, the use of only one type of mass spectrometer cannot avoid the instrumental bias involved in mass spectrometric measurements. In spite of its great success as a qualitative analysis tool, HRMS analysis has often remained semi-quantitative at best, comparing two or more samples at the same experimental conditions. As a result it always has a risk to misrepresent the actual sample composition. For example, a highly oxygenated sample with low mass components may appear to be less oxygenated if the instrument has a significant mass discrimination in the low mass ion range; this could easily happen in the FTICR if one ignores the 'time-of-flight effect'.[6]

Recently, Pomerantz and coworkers proposed the use of an orbitrap mass analyzer for routine finger printing analysis of major crude oil components and successfully demonstrated the similarity of heteroatom class distributions and double bond equivalent (DBE) distributions between FTICR and orbitrap data.[7] While this study suggests an orbitrap mass analyzer may provide sufficient mass resolution for major crude oil compounds, other differences between the two instrumentations (such as mass discriminations) were not investigated. In the current study, we are trying to address this issue in an effort to minimize possible bias in HRMS analysis for bio-oils. The three high-resolution mass spectrometers we adapted in this study are FTICR, orbitrap, and quadrupole-time-of-flight (Q-TOF). Their operation principles are quite different from each other, not only in mass analyzer but also in ionization source designs; therefore, the comparison is expected to reveal any instrument dependent bias.

Laser desorption ionization (LDI) that we used in the previous study has a few limitations in that 1) volatile compounds are mostly vaporized inside the ion source vacuum before the data acquisition, 2) only those compounds that have absorption in the laser wavelength can be desorbed and ionized, and 3) it has a technical difficulty due to the need to minimize the laser power to prevent in-situ aggregation from occurring.[5] These limitations reduced the number of compounds that can be ionized, and the reduced complexity allowed us to successfully analyze bio-oil samples even with the low mass resolution version of orbitrap. Although LDI was successful in characterizing lignin dimeric and trimeric components in bio-oils, it failed to reveal other bio-oil compounds including volatile compounds and holocellulose pyrolysis products. Electrospray ionization (ESI) has been used to examine polar compounds in petroleum crude oils. For example, negative ESI has been used to examine chemical speciation of calcium and sodium naphthenate deposits.[8] Positive and negative ESI has also been utilized to identify nonvolatile polar acidic and basic emulsion stabilizers in different types of petroleum oils.[9]

Here we adapt electrospray ionization with high resolution mass spectrometry for a petroleomic analysis of bio-oils to characterize various classes of compounds that were not possible in the previous study. Three common high resolution mass spectrometers were also compared in this study to examine instrumental artifacts and biases and to validate our methodology. In particular, we are adapting negative ion mode because most of the bio-oil compounds are protic and can be deprotonated and detected in negative ESI.

Methods

Fast Pyrolysis

The bio-oil samples were provided by Robert Brown at Iowa State University. In summary, the bio-oil used for this study was produced by fast pyrolysis of red oak with a pilot scale fluidized bed reactor located at ISU's Biocentury Research Farm. The sample was recovered from the third fraction of the bio-oil recovery and further details on this system can be found elsewhere.[10] This recovered fraction of bio-oil is dark in color, viscous and is chemically unstable; therefore, to slow chemical transformation the samples were diluted in methanol at a concentration of 1mg mL^{-1} and stored at 4^0C until analysis. Nalgene bottle were used to store bio-oils because of its chemical resistivity.

Mass Spectrometric Data Acquisition

The bio-oil was analyzed using ESI in negative ion mode. The stock solution of bio-oils was further diluted to a final concentration of 0.1 mg mL^{-1} in methanol and water (v/v 50:50). Three different high resolution mass spectrometers were adapted for the analysis: FTICR (7T Solarix; Bruker, Billerica, MA), orbitrap (LTQ-Orbitrap Discovery; Thermo Scientific, San Jose, CA) and Q-TOF (6540 Q-TOF; Agilent, Santa Clara, CA). Each instrument was carefully tuned to provide adequate signal in the mass range of interest and to minimize possible aggregations and/or fragmentations. Various ion flight times were tested for FTICR and 0.4ms was used in the final data acquisition. The data acquisition size for FTICR was 2M with a transient length of approximately 0.9 seconds and the orbitrap's transient length was approximately 0.5 seconds.

Data Analysis

The peak list was produced by DataAnalysis (Bruker) and QualBrowser (Thermo Scientific) for FTICR and orbitrap data, respectively and imported to Composer (Sierra Analytics, Modesto, CA) for the petroleomic analysis. The internal mass calibration was performed by DataAnalysis for FTICR data and by Composer for Orbitrap data based on a list of known bio-oil components that were previously identified using an internal 3-point calibration of added known compounds. The mass accuracy of Composer was limited to 3ppm in the chemical composition analysis and the relative ion abundance was limited to 0.1%.

Results and Discussion

Comparison of High Resolution Mass Spectrometers

The most significant difference we expect among the three mass spectrometers is its mass discrimination; some instrumentations and/or experimental conditions have higher sensitivity for low mass ions vs. high mass ions or vice versa. Hence, the experimental conditions should be carefully optimized to minimize mass discriminations for the mass range of interest. Additionally, ion source conditions and ion guide voltages should be optimized to reduce aggregations and minimize fragmentations. After careful optimization of experimental conditions, with model compounds and bio-oil, similar spectral patterns were obtained for the bio-oil sample with all three high-resolution mass spectrometers as shown in Figure 1. There are some minor differences among the three spectra that should be noted. First, the relative ion abundance of high

mass ions, especially those of m/z 200 or greater, is higher for FTICR compared to orbitrap or Q-TOF. In contrast, the FTICR data shows mass discrimination against very low mass ions, especially those below m/z 131. The relative abundance of high mass ions vs. low mass ions is easily affected by ion tuning parameters and we cannot decide which spectrum represents ion abundances closer to the real compositions of the bio-oil.

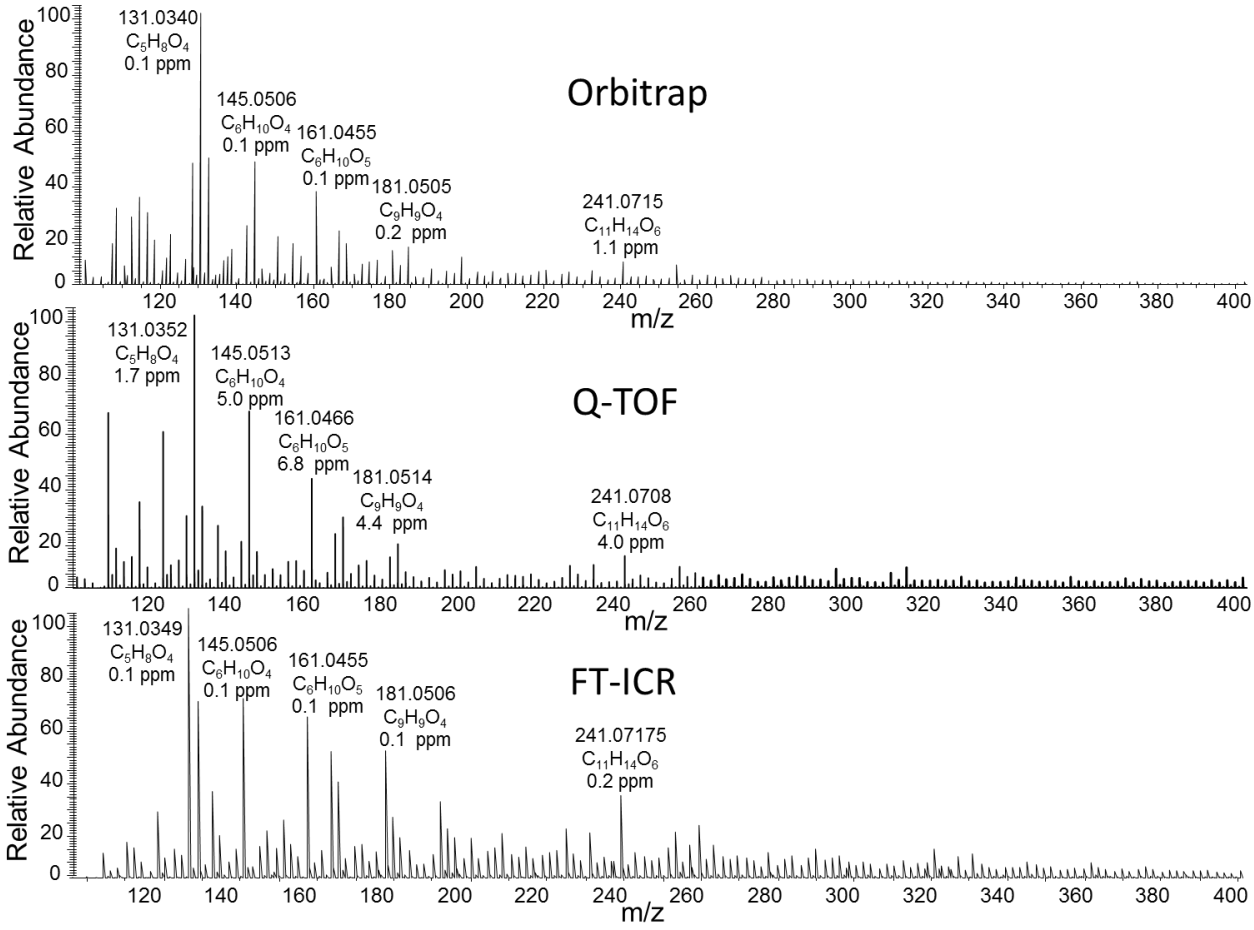


Figure 1. High-resolution mass spectra of red oak bio-oils obtained using orbitrap, Q-TOF, and FT-ICR. All of the peaks are deprotonated negative ions produce by ESI. Chemical compositions are shown for the corresponding non-deprotonated forms.

In FTICR, the ion flight time between the quadrupole and the ICR cell is the instrument parameter that most affects the transmission of high mass ions vs. low mass ions. The ion flight time (also called 'time-of-flight') determines how long the ICR cell has the gate open for the ion injection during which the trapping of injected ions occurs by the high magnetic field inside the cell. If sufficient time is not given some high mass ions might not have arrived at the cell. On the other hand, if the gate is open too long some low mass ions will be lost. The effect of the ion flight time on mass discrimination is notorious in FTICR, and typically there is no good way of efficiently trapping ions in a very wide mass range.[6] Therefore, it should be optimized depending on the mass range of interest. The effect of the ion flight time on ion transmissions was examined with lithium acetate cluster ions and bio-oil (Supplementary Figures 1 and 2). At ion flight time of 0.4ms, ion transmission is significantly reduced for high mass ions above m/z 400. However, at ion flight time of 0.6ms, ion transmission is less than half for low mass ions below m/z 200 compared to that of 0.4ms. For the data shown in Figure 1, we used an ion flight time of 0.4ms because 1) it is most similar to those of orbitrap and Q-TOF and 2) most of the bio-oil ions are below m/z 400 in negative ESI mode.

Another very important characteristic of these instruments that needs to be taken into consideration for the analysis of complex mixtures, like bio-oils, is the mass resolving power ($R = m/\Delta m$; Δm is typically defined as full-width at half maximum). The ability to resolve complex mixtures effectively is what allows for an accurate mass measurement that gives a confident assignment of the peaks. The Orbitrap Discovery used in the current study has a lower mass resolving power than FTICR. As shown in

Supplementary Figure 3, orbitrap has 5-10 times lower resolving power than that of a 7T FTICR in m/z range of 100-400. The mass resolving power of Q-TOF is even lower for the mass range of interest: 6-30 times lower compared to 7T FTICR. A new ultrahigh-resolution orbitrap, Orbitrap Elite (not used in the current study), is expected to have an equivalent mass resolving power to typical 7T FTICR; however, a newly developed FTICR cell has capability of greatly improving the mass resolving power and mass resolving power of 24,000,000 at m/z 609 is reported for 7T FTICR.[11] There are also several other data processing techniques that can be done to improve resolution of these instruments.[12-13]

Careful inspection and comparison was made for each peak between FTICR and orbitrap spectra in Figure 1 to determine whether they have sufficient mass resolution to differentiate the two closest peaks. All the peaks in the orbitrap spectrum have sufficient resolution for m/z 300 or below as confirmed with FTICR data (Supplementary Figure 4 A & C). As the m/z increases, the resolution decreases and the bio-oil molecular complexity increases; therefore, the Orbitrap Discovery does not have enough resolution to distinguish some peaks above m/z 300. For example, the peak at m/z 315.071 in orbitrap data (Supplementary Figure 4B) is split into two distinct peaks in the FTICR spectrum (Figure 2D, m/z 315.0715 and m/z 315.0740). However, the ion abundances for m/z > 300 are very low with the relative intensity less than ~1%, and those peaks with near-by interferences are even lower, <0.1%. Hence, the overall picture of orbitrap data is not affected by the interference peaks. However, based on the insufficient mass

resolution for the higher m/z ions in orbitrap, the focus of the chemical composition analysis was done with the FTICR data.

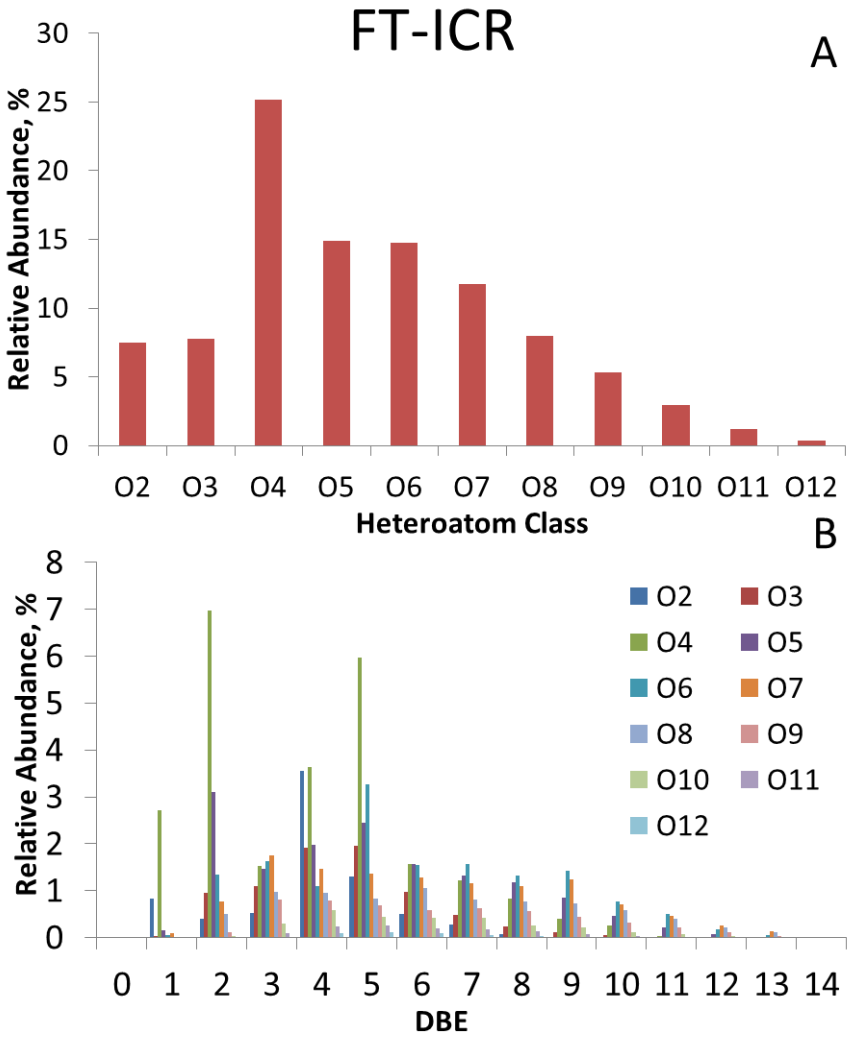


Figure 2. (A) Heteroatom class distribution and (B) DBE distribution for the data obtained with FT-ICR

Chemical Composition Analysis

Figure 2A shows the heteroatom class distribution for the data obtained from FTICR. The chemical composition analysis showed several heteroatom classes consisting of 2 to 12 oxygens and the most abundant heteroatom class having 4 oxygens

accounts for 25% of the identified peaks. The next most abundant classes of compounds are O5 and O6 heteroatom classes with approximately 15%. Double bond equivalency (DBE) distribution shows the most abundant DBE for O4 compounds is DBE of 2 followed by DBE of 5 (Figure 2B). In general, the lower oxygen heteroatom classes (O1-O6) have lower DBE values of 1 to 6 whereas the higher oxygen heteroatom classes (O7 or higher) have higher DBE values of 6 to 12. Figure 3 shows the contour plots of carbon number vs. DBE for O3-O10 compounds. An increase in DBE and carbon number is accompanied by higher oxygen content. Further in-depth discussion is made in the following two sections about the contour plots. Similar chemical composition results were obtained from the data set with orbitrap data (Supplementary Figures 5-7) except for some minor differences mostly coming from the differences in mass discrimination.

Most of the major compounds shown in Figure 1 are expected to be volatile considering their low m/z values; therefore, the analysis could be possible in GC-MS as long as they can be separated in GC and identified by EI-MS spectra. Some peaks are in fact matching with previously known compounds in GC-MS of bio-oils. Supplementary Table 1 shows a summary of known GC-MS compounds in the literature[14, 21-24] classified by their heteroatom, DBE and carbon number; thus, each compound in the table can be easily matched to each dot in Figure 3. For instance, all the O3 compounds in the table is showing up as major dots in Figure 3. It demonstrates that many volatile compounds currently known in GC-MS can be characterized in negative ESI.

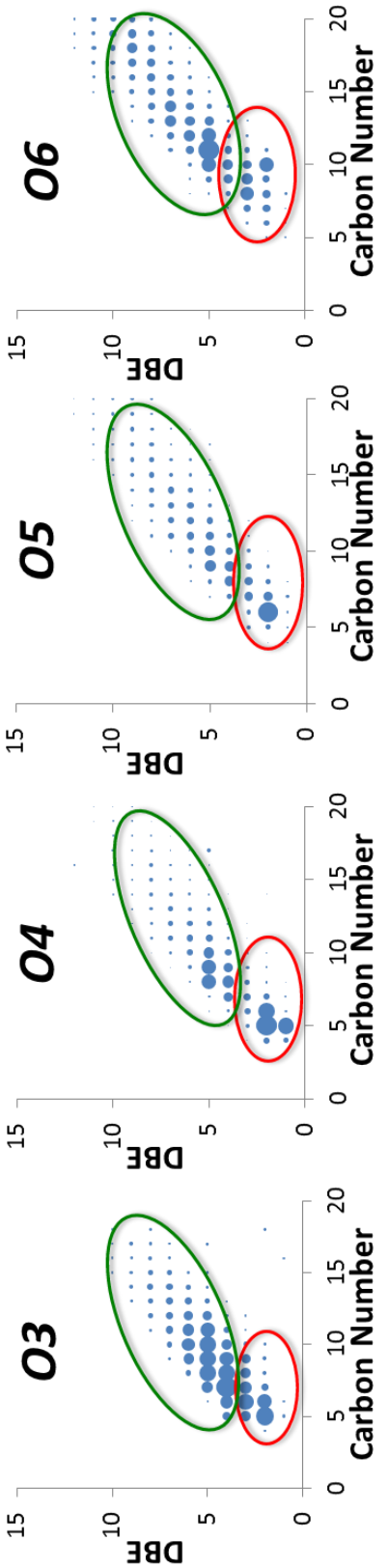


Figure 3. Contour plots for O3-O10 heteroatom classes with regards to carbon number versus DBE. Red and green circles represent sugaric and phenolic compounds.

Negative ion mode seems to be especially effective considering most of the bio-oil compounds contain hydroxyl (-OH) or carboxylic (-COOH) groups that can be deprotonated. Low molecular weight (MW<~100) or non-polar compounds are missing in this analysis; however, we could detect non-volatile or partially volatile compounds instead. The number of dots in Figure 3 is much more than the compounds listed in Supplementary Table 1. It suggests the molecular complexity of bio-oils is much more immense than has been analyzed by typical GC-MS analysis. According to our Kendrick mass defect analysis, we have identified over 800 chemical compositions and only about forty of them are known in GC-MS. This is not surprising considering the molecular complexity of biomass materials, the various pyrolysis reactions possible, and the secondary/aging reactions that are continuously happening during the condensation process and on the shelf.

When Supplementary Table 1 is compared with Figure 3, one can notice not only non-volatile compounds, but also many volatile compounds are missing in GC-MS analysis; e.g., most of small dots with low carbon and DBE numbers in Figure 3. The shortcoming of GC-MS analysis is mostly attributed to 1) insufficient chromatographic separation in typical GC-MS analysis and 2) unavailability of many compounds in EI-MS database. The former can be greatly improved by recent GCxGC technology.[15-16] Unfortunately, there is no good solution that can improve chemical identifications with the current EI-MS approach. This is because many of the pyrolysis products are occurring through multi-step radical reactions, and it is impossible to individually synthesize and purify each compound to add it into the EI-MS database. Some major

pyrolysis products are stable and previously known, but many others still remain unknown.

Holocellulose pyrolysis products or 'sugaric' compounds

One can note in Supplementary Table 1 that most of the bio-oil compounds are derivatives of phenol, guaiacol and syringol; these are well known lignin pyrolysis products that originated from the monomers of coumaryl, coniferyl and sinapyl alcohol. We call these compounds 'phenolic compounds'. However, there are only a handful of cellulose or hemicellulose pyrolysis products analyzed by GC-MS; e.g., levoglucosan (1,6-anhydro-β-D-glucopyranose; the most known cellulose pyrolysis product), levoglucosenone, dianhydroglucopyranose, anhydropentose (such as anhydroxylpyranose or anhydroarabinofuranose), and furfural and its derivatives (mostly m/z <100 not included in Supplementary Table 1). The lack of known holocellulose pyrolysis products might be mostly attributed to 1) less complexity of cellulose and hemicellulose compared to lignin and 2) difficulty of characterizing EI-MS spectra of these polyhydroxyl compounds. The latter is quite problematic in the analysis of holocellulose pyrolysis products using GC-MS. For example, EI-MS spectra of various hexoses (e.g., glucose, galactose, and mannose) look almost identical to each other. In addition, because of the extensive fragmentation in typical electron ionization condition, EI-MS spectra of the hexoses look very similar to that of pentoses such as ribose and arabinose. Derivatization or chemical ionization (CI) could be utilized to minimize the fragmentation but EI-MS spectra of derivatized compounds or CI-MS spectra are not available for most of the compounds.

We examined our data closely to determine whether our technique could identify these holocellulose pyrolysis products. Especially since there is virtually no fragmentation in ESI, it is expected to be very useful in the analysis of holocellulose pyrolysis products.[17] As described below, we found many holocellulose pyrolysis products corresponding to poly-hydroxyl cyclic hydrocarbons, such as levoglucosan and/or its derivatives. We call these compounds 'sugaric compounds' which are meant to describe 'derivatives of sugars' in opposition to 'phenolic compounds'. They cannot be called oligosaccharides or carbohydrates because many of them do not have the chemical formulae of $C_n(H_2O)_m$. Sugaric compounds are distinguishable from phenolic compounds by low unsaturation (or double bond equivalence (DBE)) and high oxygenation. Namely, phenolic compounds should have DBE value of at least 4 while levoglucosan has a DBE value of 2. A group of O4 and O5 class compounds with low DBE values (DBE <4) in Figure 3 are clearly distinguished from phenolic compounds (DBE $\geq$ 4) and are suspected to be sugaric compounds.

Figure 4 shows a graphical representation of some of those sugaric compounds. The O5 DBE2 with 6 carbons (C6) corresponds to the chemical composition of $C_6H_{10}O_5$ which matches that of levoglucosan. The O4 DBE2 C5 corresponds to the chemical composition of $C_5H_8O_4$ which matches that of anhydropentose; it could have resulted from either anhydrolysis of pentose or CH_2O-loss from levoglucosan during pyrolysis. There is no compound known in GC-MS of bio-oils for $C_5H_{10}O_4$ (O4 DBE1 C5) or $C_6H_{10}O_4$ (O4 DBE2 C6), but possible structures include deoxypentopyranose and glucal as suggested in Figure 5. One can also find minor spots around these major compounds

in Figure 4, such as $C_4H_6O_4$ (O4 DBE2 C4), $C_7H_{12}O_4$ (O4 DBE2 C7), $C_4H_8O_4$ (O4 DBE1 C4), and $C_6H_{12}O_4$ (O4 DBE1 C6), illustrating the molecular complexity of holocellulose pyrolysis products.

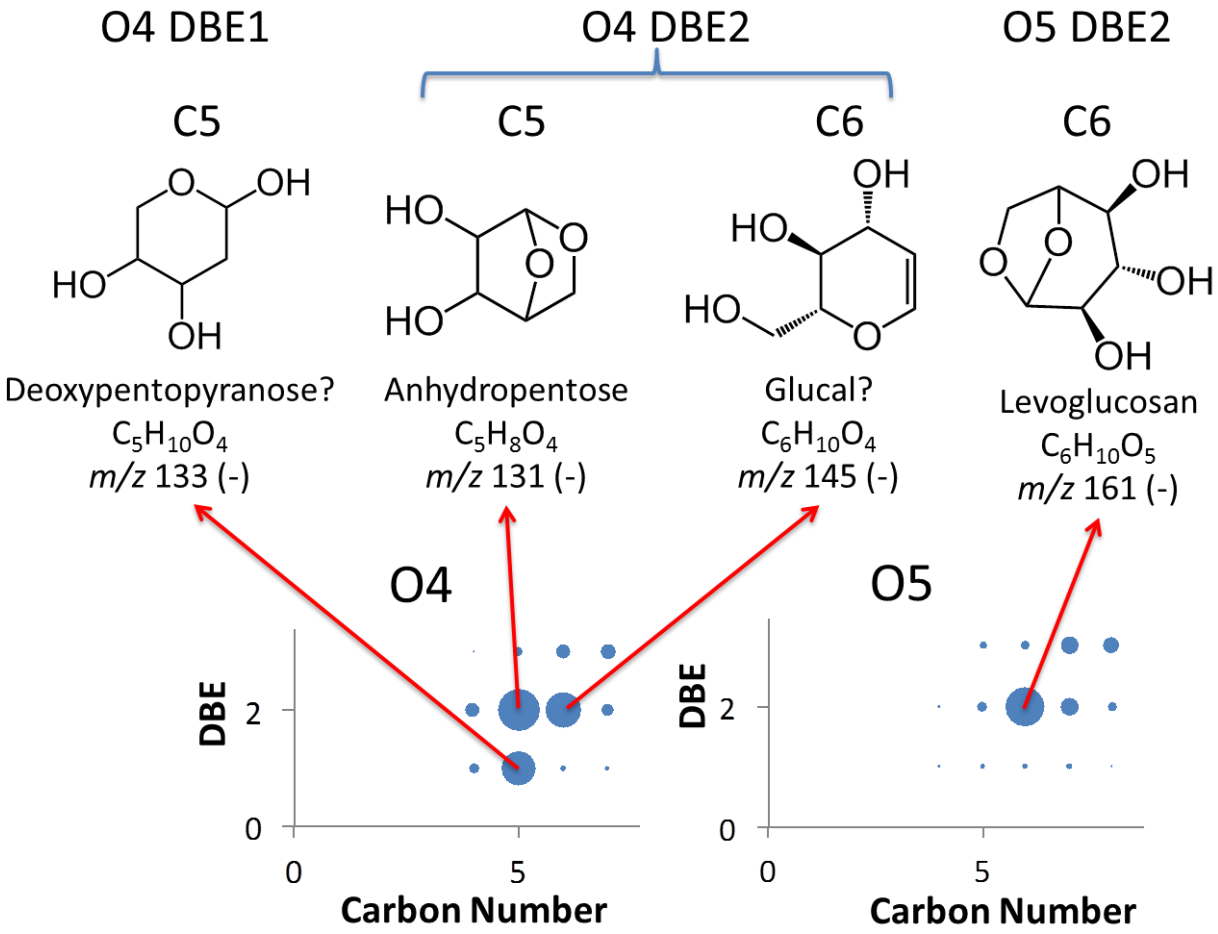


Figure 4. Illustration of possible chemical structures for a few pyrolysis products of holocellulose.

The next question we had was how sugaric vs. phenolic compounds are changing as the number of oxygens increase. As illustrated in Figure 4, we believe those compounds with DBE less than 4 are mostly sugaric compounds while those with DBE of 4 or higher are phenolic. As circled in Figure 3, sugaric (red) and phenolic (green)

compounds are separated from each other although the distinction is not as clear at low (O3) or high (O6 or higher) oxygen class as in O4 and O5 compounds. The highest abundance ion in sugaric compounds is gradually moving from low carbon number to high carbon number as oxygen number increases: C5 for O3 and O4 →C6 for O5 →C8 (or C10) for O6 →C11 for O7 →C12 for O8 →C13 for O9 →C14 for O10. The DBE value also changes from 2 (for O3-O5) to 3 (for O6-O9) and then to 4 (for O10).

For O6 or higher oxygen classes, some sugaric compounds might have DBE values of 4 and overlap with phenolic compounds is inevitable (see the overlap of two circles). Those compounds on the borderline (O6 or higher with DBE values of 4) might be difficult to determine if they are sugaric or phenolic or a combination of both. The increase of unsaturation at high oxygen class might be related with the dimerization of sugaric compounds. For example, cellobiosan ($C_{12}H_{20}O_{10}$; a known sugar pyrolysis product) is a heterodimer between glucose and levoglucosan and has a DBE of 3.[18-19]A spot on the contour map corresponding to cellobiosan (O10 DBE3 C12) can be found in the contour plot for O10 but in much less intensity than its alkyl derivatives (C13 or higher). This suggests the occurrence of various chemical reactions during and/or after pyrolysis. An in-depth understanding beyond this is difficult because of the uncertainty of transmission/ionization efficiency for each ion species and the difficulty of quantifying the results. This will be further discussed in the next section.

pH Dependence of Various Bio-oil Compounds

While overall spectral patterns are similar between instruments in Figure 1, relative intensities of some peaks are significantly different; e.g., m/z 129 and 143. Even for the same instrument we found that relative intensities of some peaks are easily affected by various experimental conditions including sample aging. It is not surprising to see some changes of the bio-oil components as it ages, but often the difference appears to be more than just aging, especially in negative ESI mode. For example, the mass spectra in Supplemental Figure 2 were obtained about one year after the spectra in Figure 1 was obtained. The relative peak height for m/z 161 (levoglucosan) is twice as high in Supplemental Figure 2 as that of Figure 1. Therefore, we performed a simple experiment to determine how the relative abundances of bio-oil components are affected by other variables, especially pH because it is known that pH decreases as bio-oil ages. Figure 5 shows the FTICR spectra obtained at four different pHs: pH of 3.5, 5, 7, and 9. Acetic acid and ammonium hydroxide were used to adjust pH. One can notice significant difference between acidic pH and basic pH. At low pH values; m/z 131, 145, 161, 169, and 241 are the major components, but at high pH values; m/z 167, 181, and 219 are the major components. This difference partially explains the variability in data acquisition and the minor difference between instruments. Bio-oil is a very complex mixture with various acidic molecules that are reactive, and the relative ion abundances of some of its components can be easily affected in negative ESI mode by a simple pH change or by other modifiers such as ammonium formate (data not shown).

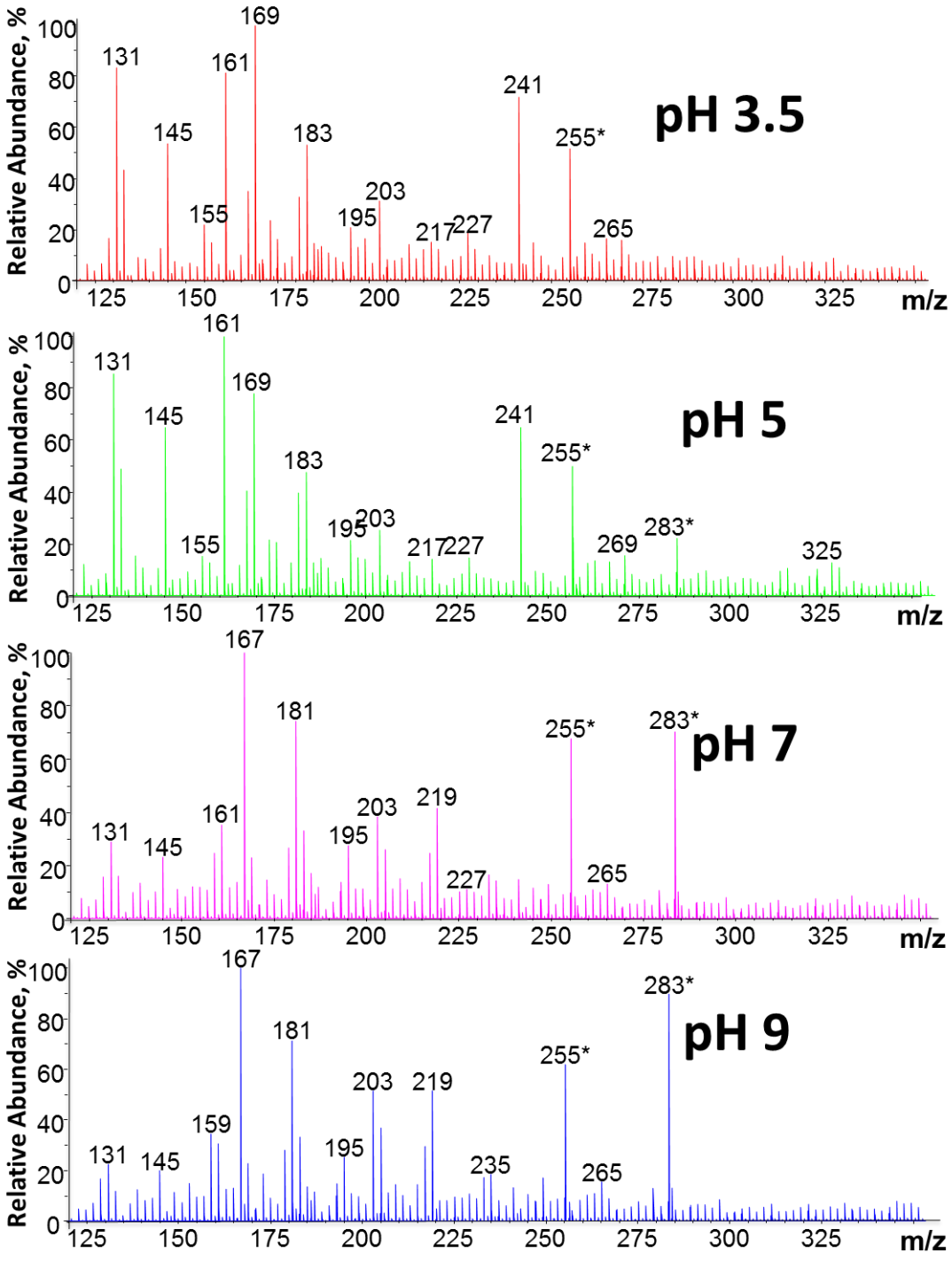


Figure 5. pH dependence of negative ESI mass spectra of bio-oils obtained with FT-ICR (*) Contamination.

Table 1. pH dependence of a few major peaks in bio-oils.[a]

m/z	Compound Classifications	pH 3.5	pH 5	pH 7	pH 9
131.0350	O4 DBE2	90%	90%	30%	25%
145.0506	O4 DBE2	55%	70%	25%	20%
161.0455	O5 DBE2	80%	100%	30%	30%
167.0350	O4 DBE5	35%	40%	100%	100%
169.0506	O4 DBE4	100%	80%	25%	25%
181.0506	O4 DBE5	35%	40%	80%	80%
183.0663	O4 DBE4	55%	50%	35%	35%
219.0663	O4 DBE7	5%	5%	40%	50%
241.0717	O6 DBE5	70%	70%	15%	15%

a. Relative ion abundance in each spectrum shown in Figure 5.

Table 1 summarizes the pH dependence of a few major compounds. There are a few important findings in this table. First, ion signals for sugaric compounds (O4/O5 DBE2) are greatly suppressed at high pH. This is contradictory to our initial thoughts that basic conditions may improve ion signals for high pKa compounds, such as levoglucosan, and that acidic condition may suppress deprotonation of these compounds. However, our data suggests that moderate acidic condition (pH5) gives the best ion signals for sugaric compounds in negative ESI. A similar effect has been reported by Wu and co-workers where an increase of ion signals for androgen receptor modulators was seen in negative ESI by adding acetic acid.[20] They attributed it to the effect of counter anions (i.e., acetate) in ESI droplets. This same explanation could be possible for sugaric compounds. Namely, in weakly acidic solutions, these molecules are not deprotonated because the pH is less than the pKa. However, in negative ESI condition, most protons are electrochemically reduced to hydrogen at the spraying electrode and acetate anions are accumulated in the ESI droplets. Acetate anions have a high proton

affinity and can easily abstract protons from sugaric compounds and generate deprotonated ions.

A second finding in Table 1 is some phenolic compounds have better ion signals at low pH (e.g., m/z 169, 183, and 241) while other phenolic compounds have better ion signals at high pH (e.g., m/z 167, 181, and 219). We currently do not have a good understanding of this behavior especially since we do not know their exact chemical moieties; however, we noticed the overall ion signals are greatly suppressed at high pH. It is presumably due to ion suppression caused by ammonium hydroxide that was added to increase the pH. If some compounds are subjected to less ion suppression than others, then their relative ion signals would appear to be better. Further study will lead to a better understanding of the ionization behavior of these compounds and better characterization of some of the unknown compounds.

Conclusions

We successfully demonstrated electrospray ionization in negative ion mode combined with HRMS-based chemical composition analysis as an efficient way to characterize bio-oils. Specifically, most of the bio-oil compounds have hydroxyl functional groups that can be deprotonated and analyzed in negative ESI. According to our analysis, the complexity of bio-oils is much more immense than typically understood by GC-MS with over eight hundred chemical compositions assigned in our analysis. The majority of compounds are low mass ions and they are expected to have some volatility. In contrast, only about forty chemical compositions were previously known in

GC-MS. There are a few classes of compounds missing in the current study; aprotic and/or non-polar compounds will not be analyzed in negative ESI and non-volatile less-polar compounds may not be as efficiently ionized as others. Considering the missing compounds in negative ESI and structural isomers possible for each chemical composition, we suspect the actual complexity of bio-oils may be well over a thousand compounds.

In the development of a reliable HRMS-based petroleomic analysis using negative ESI, we have performed a comparative study between three common high-resolution mass spectrometers. Although each instrument has its own limitations in mass spectrometric measurement, we could obtain almost identical spectra assuring our analysis. We found that FTICR analysis should be performed carefully because it has significant low mass discriminations, especially arising from the ion flight time. The other two high-resolution mass spectrometers adapted in the current study (Orbitrap Discovery and Q-TOF) have much lower mass resolution, but were still useful for confirmation of the FTICR data. Most of the known bio-oil compounds in GC-MS with the mass over 100 are detected in our analysis with their corresponding chemical compositions (both lignin and holocellulose pyrolysis products). Phenolic and sugaric compounds could be distinguished in the contour plot of DBE vs. the number of carbons based on their difference in DBE values.

The current approach has a few shortcomings compared to GC-MS analysis. Most of all, quantitative analysis is expected to be a daunting task especially because negative ESI is easily affected by several experimental parameters as confirmed by the

pH dependence experiment. Second, many of the chemical compositions have structural isomers and cannot be distinguished with accurate mass only. MS/MS might distinguish some structural isomers but most of them have several near-by peaks and the effective isolation of precursor ions may be difficult. Finally, negative ESI cannot ionize non-polar and/or aprotic compounds. In addition, as an atmospheric pressure ionization method, ESI has limitation for very low mass ions; i.e. m/z <100. The future work will be focused on a better understanding of the pH and/or organic modifier effect in electrospray ionization and to establish quantitative methodology. We are currently working on HRMS analysis of bio-oils using other ionization techniques (such as atmospheric pressure photoionization (APPI) and atmospheric pressure chemical ionization (APCI)) to have a comprehensive understanding of bio-oils especially for non-volatile compounds that are not amenable to GC-MS analysis and non-polar compounds that are not amendable to ESI.

Acknowledgements

This work was supported by a grant from ConocoPhillips. The authors thank Robert C. Brown, Center for Sustainable Environmental Technology at Iowa State University, and his group members for bio-oil samples and valuable discussions. The authors are also grateful to David Stranz, Sierra Analytics, for kindly providing an evaluator version of Composer software for this study. E.A.S. acknowledges partial support from the Graduate Assistance in Areas of National Need (GAANN) fellowship from the U.S. Department of Education. S.P. was partially supported by a fund from

National Research Foundation of Korea and the U.S. Department of Energy (DOE), Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences. The Ames Laboratory is operated by Iowa State University under DOE Contract DEAC02-07CH11358. The FTICR used in the current study was acquired through a National Science Foundation Major Research Instrument grant to the W.M. Keck Metabolomics Research Facility at ISU.

References

(1) Gellerstedt, G.; Li, J.; Eide, I.; Kleinert, M.; Barth, T. *Energy Fuels* **2008**, 22, 4240-4244.

(2) Kleinert, M.; Barth, T. *Energy Fuels* **2008**, 22, 1371-1379.

(3) Oasmaa, A.; Czernik, S. *Energy Fuels* **1999**, 13, 914-921.

(4) Marshall, A.; Rodgers, R. P. *Acc. Chem. Res.* **2004**, 37, 53-59.

(5) Smith, E.; Lee, Y.J. *Energy Fuels* **2010**, 24, 5190-5198.

(6) Barrow, M.; McDonnel, L.; Feng, X.; Walker, J.; Derrick,P. *Anal. Chem*. **2003**, 75, 860-866.

(7) Pomerantz, A.; Mullins, O.; Paul, G.; Sanders,M. *Energy Fuels* **2011**, 25, 3077-3082.

(8) Mapoleo, M.; Standford, L.; Rodgers, R.; Yen, A.; Debord, J.; Asomaning, S. Marshall, A. *Energy Fuels* **2009**, 23, 349-355.

(9) Hapala, M.; Purcell, J.; Saarela, V.; Franssila, S.; Rodgers, R.; Hendrickson, C.; Kotiah
Standford, L.; Rodgers, R.; Marshall, A. *Energy Fuels* **2007**, 21, 973-981.

(10) Pollard, A.S.; Rover, M.R.; Brown, R.C. *J. Anal. Appl. Pyrolysis* **2012**, 93, 129-138.

(11) Nikolaev, E.; Boldin, I.; Jertz, R.; Baykut, G. *J.Am.Soc.Mass Spectrom.* **2011**, 22, 1125-1133.

(12) Aarstol, M.; Comisarow, M. *Int. J. Mass Spectrom. Ion Processes* **1987,**76, 287-297.

(13) Qi, Y.; Thompson, C.; Van Orden, S.; O'Connor, P. *J. Am. Soc. Mass Spectrom.* **2011**, 22, 138-147.

(14) Mullen, C.; Boateng, A. *Energy Fuels* **2008,** 22, 2104-2109.

(15) Ozel, M.; Hamilton, J.; Lewis, A. *Envir Sci & Tech.* **2011**, 45, 1497-1505.

(16) Marsman, J.H.; Wildschut, J.; Mahfud, F.; Heeres, H.J. *J. Chromatogra. A.* **2007,** 1150, 21-27.

(17) Gambaro, A.; Zangrando, R.; Gabrielli, P.; Barbante, C.; Cescon, P. *Anal. Chem.* **2008,** 80, 1649-1655.

(18) Fraser, M.P.; Lakshmanan, K. *Environ.Sci. Technol.* **2000**, 34, 4560-4564.

(19) Mosier, N.S.; Ladisch, C.M.; Ladisch, M.R. *Biotechnol. Bieoeng.* **2002**, 79, 610-618.

(20) Wu, Z.; Gao, W.; Phelps, M.; Miler, D.; Dalton, J. *Anal. Chem.* **2004,**76, 839-847.

(21) Bridgwater, A.V. Ed. *Fast Pyrolysis of Biomass: A Handbook*, CPL Press, Berks, UK, 2005, Vol. 2 and Vol. 3.

(22) Garcia-Perez, M.; Chaala, A.; Pakdel, H.; Kretschmer, D.; Roy, C. *Biomass and Bioenergy* **2007**, 31, 222–242.

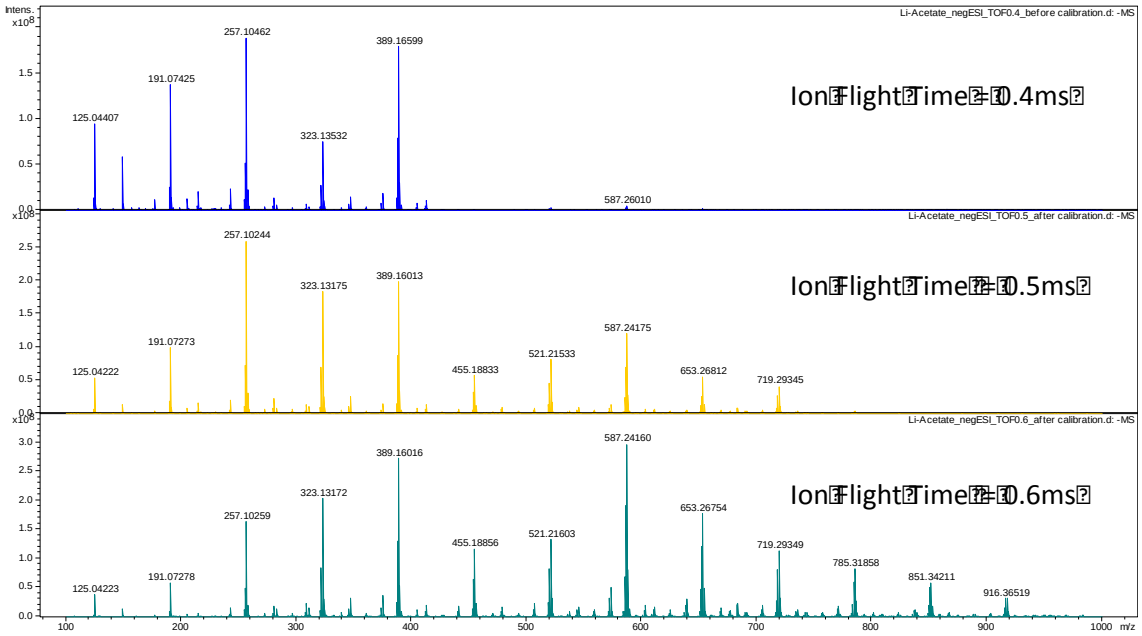
(23) Jiang, G.; Nowakowski, D.J.; Bridgwater, A.V. *Energy Fuels* **2010**, 24, 4470–4475.

(24) Rodrigues, J.; Meier, D.; Faix, O.; Pereira, H., *J. Anal. Appl. Pyrolysis* **1999**, 48, 121–128.

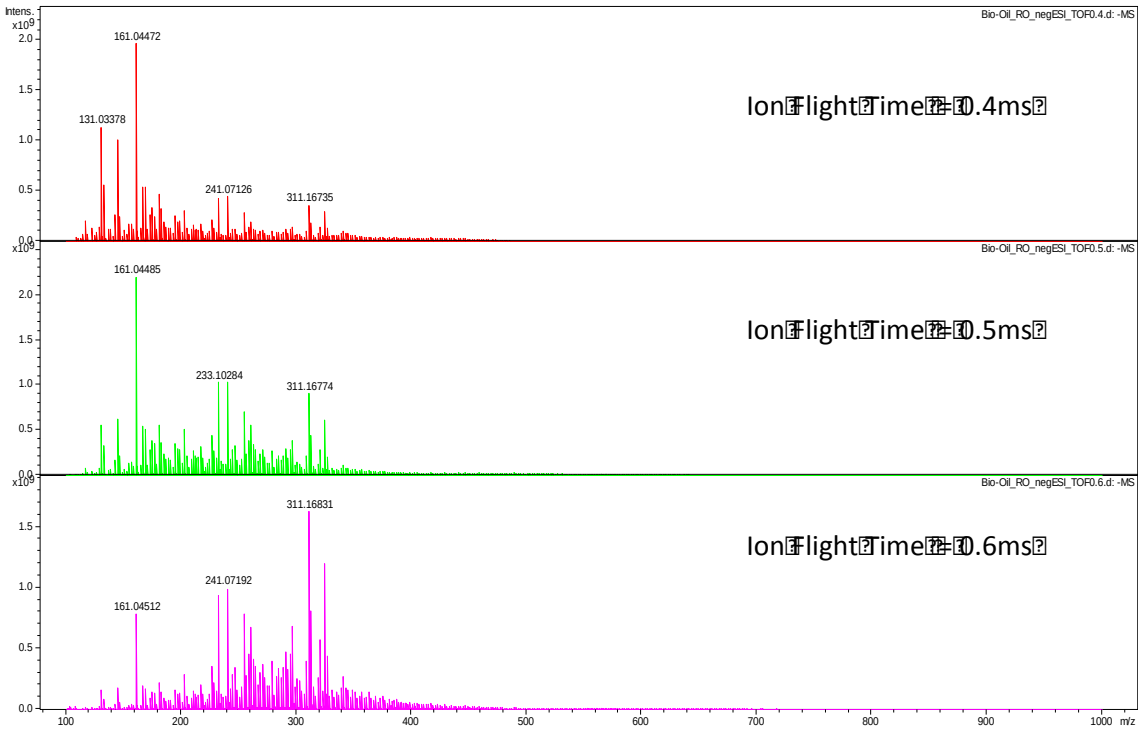
Supplementary Table 1. Classification of known bio-oil compounds in GC-MS.[a]

Heteroatom	DBE	Representative compounds in each carbon group[b]
O2	4	Furfural (C5), Benzenediol(C6), Guaiacol (C7), Methylguaiacol (C8), Ethylguaiacol(C9), Propylguaiacol (C10)
	5	Benzoic acid (C7), Hydroxymethylbenzaldehyde (C8), Vinylguaiacol(C9), Eugenol(C10)
O3	2	Oxobutanoic acid(C4), Levulic acid(C5)
	3	Hydroxymethylfurfural (C6)
	4	Levogucosenone (C6), Methoxycatechol (C7), Syringol (C8), Methyl syringol (C9), Dihydro-conferyl alcohol (C10), Propyl syringol (C11)
	5	Vanillin(C8), Veratraldehyde (C9), Vinylsyringol (C10), Allylsyringol (C11)
	6	Coniferylaldehyde (C10), hydroxydimethoxyindene (C11)
	7	2,3-dihydroxy-1H-indene-1-one (C9)
O4	2	1,5-anhydroarabinofuranose (C5)
	3	2,3-anhydro-d-mannosan (C6)
	4	Dihydrosinapylalcohol (C11)
	5	Vanillic acid (C8), Syringaldehyde (C9), Acetosyringone (C10), Syringylacetone (C11)
	6	Ferullic acid (C10), Sinapaldehyde (C11)
	10	Medicarpin (C16)
O5	1	Xylose (C5)
	2	Levoglucosan (C6)
	5	Syringic acid (C9)
O6	1	Glucose (C6)

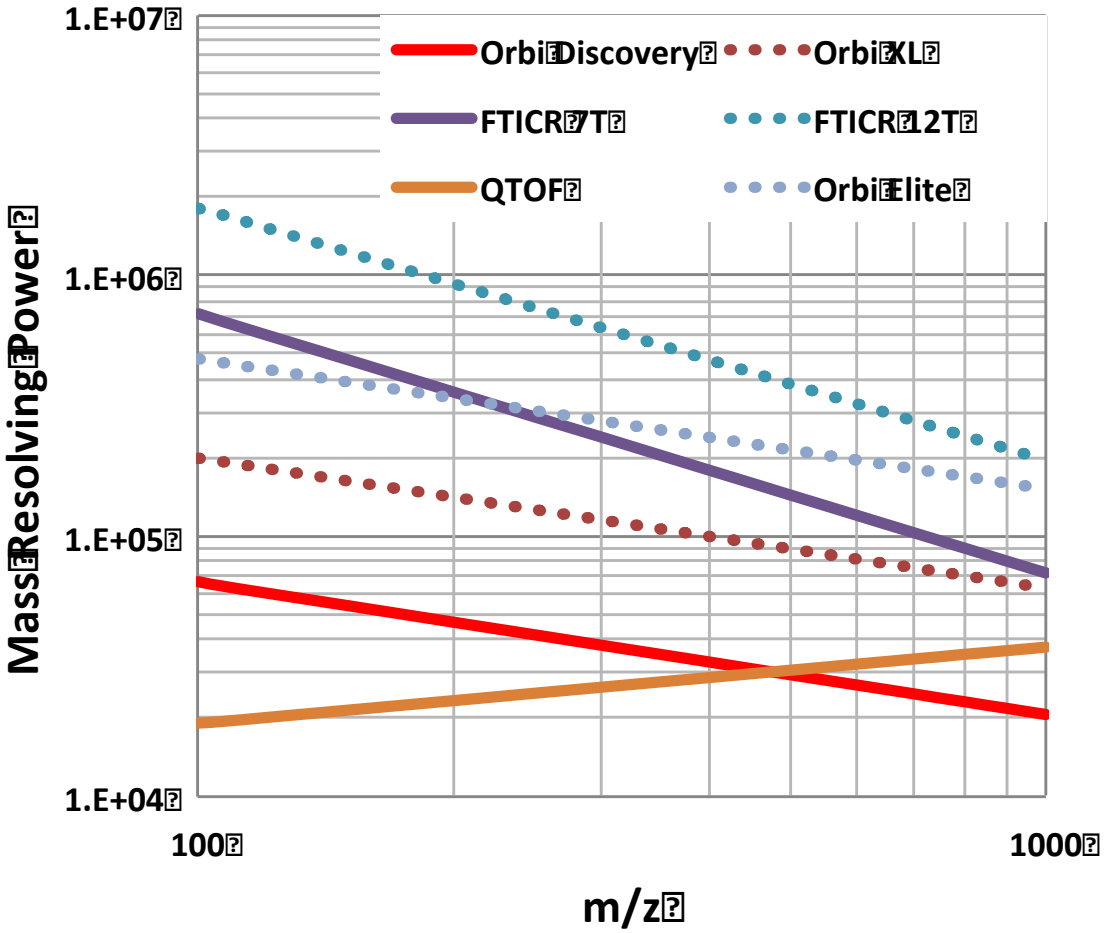
a. Bio-oil compounds known in GC-MS analysis that are detectable by the current analysis (O2 or higher oxygen compounds with m/z > 100) are classified as their oxygen, DBE, and carbon groups.
b. Only one representative compound is shown among the structural isomers with its carbon number in the parenthesis. References: 14, 21-24.



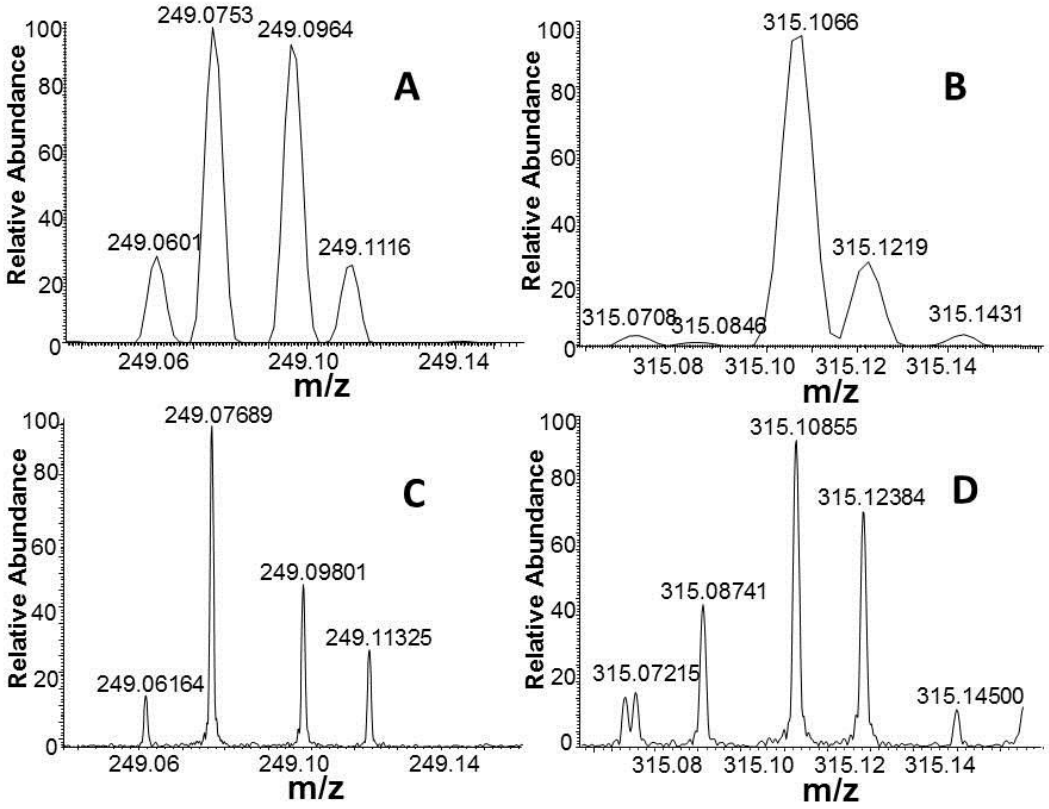
Supplementary Figure 1. FTICR mass spectra of lithium acetate cluster ions with various ion flight times (0.4ms, 0.5ms, and 0.6ms).



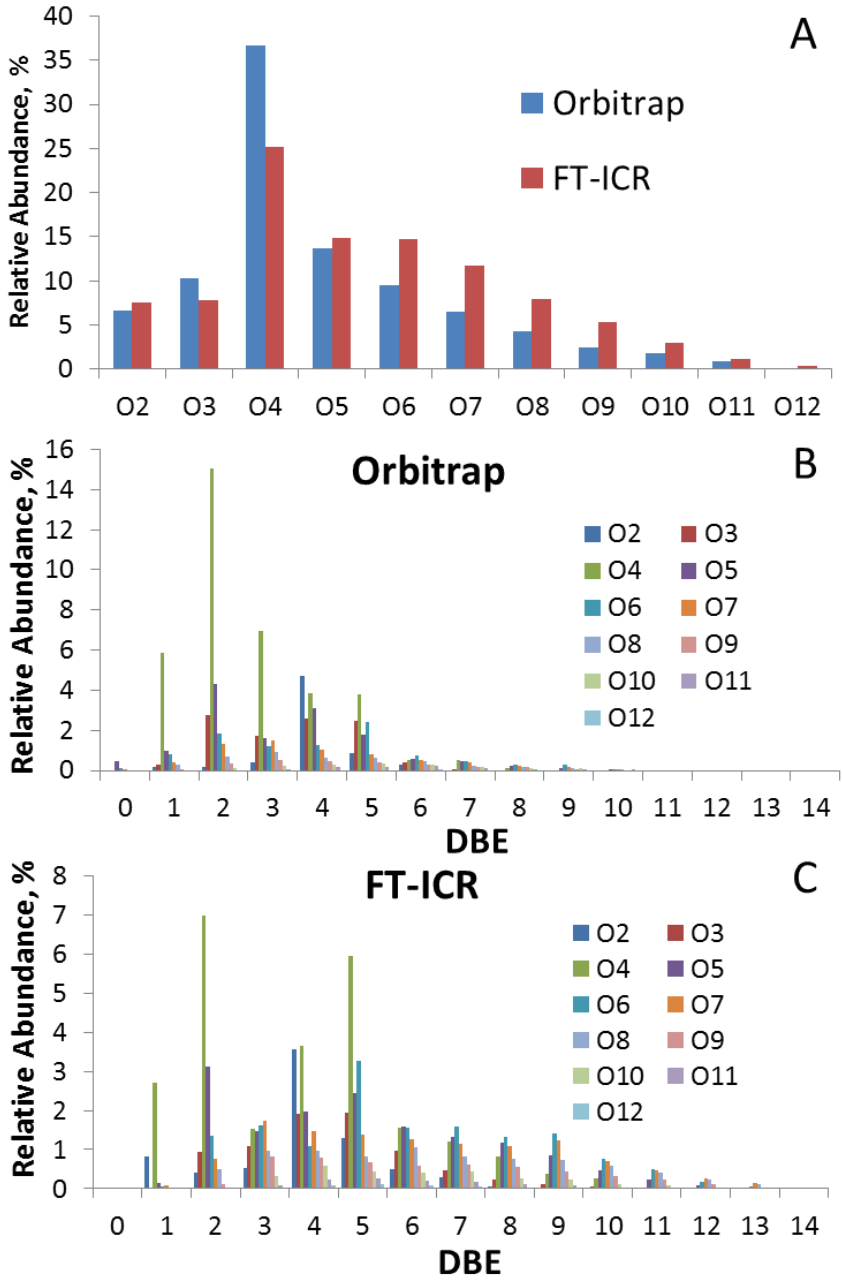
Supplementary Figure 2. FTICR mass spectra of bio-oils with various ion flight times (0.4ms, 0.5ms, and 0.6ms)



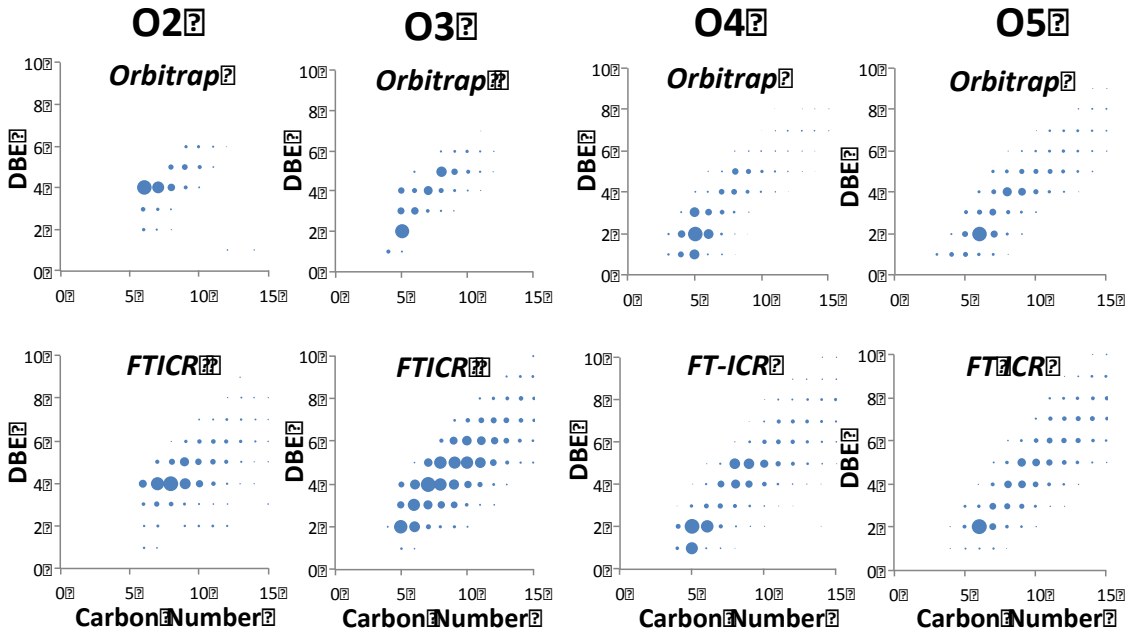
Supplementary Figure 3. Mass resolving power of a few high-resolution mass spectrometers, plotted as m/z. Q-TOF, Orbitrap Discovery, and 7T FTICR data were obtained from the current study, and those of Orbitrap XL, Orbitrap Elite, and 12T FTICR are estimated.



Supplementary Figure 4. A closer look at the resolution shows that the orbitrap (A) has sufficient resolution for lower mass ions that was confirmed with FTICR data (C). However, when looking a slighter higher mass, *m/z* of 315, the orbitrap is unable to resolve some of the minor peaks (B) as seen in FTICR data (D). The orbitrap spectrum shows a wide peak at 315.0708, but FTICR data reveals that there are two distinct peaks present and reveals a second unique peak at 315.07215.



Supplementary Figure 5. A heteroatom class distribution comparison between orbitrap and FTICR (A) shows overall similar results between the two instruments with a few exceptions that can be contributed towards mass discrimination in the FTICR for higher masses and orbitrap's mass discrimination toward lower masses. This is further confirmed in DBE distribution charts of orbitrap (B) and FTICR (C) where the major differences are seen in higher DBE values, with higher oxygen content that can be contributed with higher masses.



Supplementary Figure 6. A closer look between Orbitrap (top row) and FTICR (bottom row) for the carbon number vs. DBE of lower oxygen heteroatom classes (O2-O5). The size of the circle correlates with the intensity of the specific compound.

CHAPTER IV

STITCHING HRMS SPECTRA TOGETHER FOR A MORE COMPREHENSIVE PICTURE OF BIO-OILS

Erica A. Smith and Young-Jin Lee*

Abstract

A novel methodology for bio-oil analysis that stitches together HRMS spectra from different mass spectrometers to overcome instrument-specific mass-discrimination and resolution limitations is presented. FTICR instruments are known to suffer from low-mass discrimination, especially below *m/z* 200. Orbitrap instruments suffer from minimal low-mass discrimination, but offer limited mass resolving power relative to FTICR platforms and are susceptible to high-mass discrimination caused by space-charge effects. The disparity in resolving powers is even more pronounced for lower-end orbitrap instruments (Orbitrap Discovery). This novel stitching combines orbitrap data for low-mass ions (*m/z* <200) and FTICR data for high mass ions (*m/z*>200) into a single mass spectrum. This minimizes the effects of mass discrimination, while still providing the resolution required for analysis of peaks in the higher-mass region of the spectrum. The resulting 'stitched spectrum' was compared with GPC data of bio-oil and showed that the stitched spectrum was a better representation of the bio-oil than the unstitched spectra obtained from either instrument individually.

Introduction

Biomass fast pyrolysis is a promising technique for biorenewable energy production that converts biomass into syngas, char, and bio-oil. Bio-oil can be catalytically upgraded into biofuels or other petroleum-based products.[1] Bio-oil is a complex mixture comprised of aqueous and non-aqueous phases, both of which contain volatile and non-volatile compounds.[2] The complexity of bio-oil prevents proper characterization of all compounds with traditional analytical techniques like GC-MS. The lack of a molecular-level characterization of bio-oil is a bottleneck for process development and catalytic upgrading, thus far limiting the utility of bio-oil as an energy source.

We have previously demonstrated the use of high resolution mass spectrometry for molecular-level characterization of bio-oils.[3-5] High resolution mass spectrometry has also been widely used in petroleomics for its ability to resolve highly complex mixtures.[6, 7] Most petroleum crude oil analysis has relied exclusively on Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry (FTICR MS) due to its superior resolving power and mass accuracy compared to other high resolution mass spectrometers, specifically time-of-flight (ToF) and orbitrap instruments. However, it has been shown that for certain ionization techniques and specifically for lower mass ranges, orbitrap instruments offer sufficient resolution to effectively analyze bio-oils.[3, 4]

It is important to note that various ionization techniques produce different types of ions; for example, atmospheric pressure photoionization (APPI) produces both protonated and radical ions. The presence of both protonated and radical ions increases

the complexity of the spectrum and demands higher resolving power to differentiate the resulting peaks. We have found that the increased complexity introduced by ionization sources like APPI can limit the utility of lower-end high resolution mass spectrometers (ToF and Orbitrap Discovery) for bio-oils and petroleum crude oil. Zhurov et al. demonstrated that a higher-end orbitrap, Orbitrap Elite, has sufficient resolution for a petroleomic profile of petroleum crude oil.[8] The Orbitrap Elite offers a nominal resolving power in excess of 200,000 at *m/z* 400, far greater than the resolving power of the Orbitrap Discovery (30,000 at *m/z* 400) that was used in our previous HRMS bio-oil analysis.

A comparative study of three common HRMS instruments (ToF, orbitrap, FTICR) was performed with bio-oil to investigate the differences in resolution and mass discrimination.[4] In this prior study, each instrument was carefully tuned to minimize mass discrimination. The ion source conditions and ion optic voltages were also optimized to reduce aggregation and minimize fragmentation. This study showed that with negative ESI it is possible to obtain very similar profiles of bio-oil despite the limitations of each instrument. However, preliminary experiments with other ionization techniques (APCI and APPI) have shown that mass discrimination is more problematic with these ionization techniques and bio-oil.

The origin of these mass discrimination effects arises from the need to collect and store ions before injection into the mass analyzer. Both orbitrap and FTICR mass analyzers require on the order of 1s for data acquisition, transfer/storage, and processing. Most modern ionization techniques produce a steady, continuous flow of ions; therefore

ions must be accumulated during an ongoing scan and then pulsed into the analyzer for subsequent mass analysis and this transfer/storage leads to the loss of low-mass ions.[9] Efforts have been made to address low-mass discrimination in FTICR instrumentation; these efforts have focused on minimizing the distance between the external ion accumulation region and the center of the high resolution mass analyzer in order to prevent the loss of low-mass ions during transport.[10] Orbitrap instruments do not suffer from low-mass discrimination because its external ion accumulation source (C-trap) can be placed very close to the mass analyzer, minimizing the ion-flight time and thus low-mass discrimination.[9] However, they are susceptible to high-mass discrimination due to space-charge effects that can occur in the C-trap.[11] To reduce space-charge effects and mass discrimination, the ion population in the C-trap needs to be limited.

Unfortunately, even with modifications to instrumentation and optimization of instrument parameters mass discrimination cannot be completely eliminated. Therefore, it is beneficial to be able to 'stitch' together spectra from different instruments to overcome the limitations of the individual instrumentation being used – namely mass discrimination in both orbitrap and FTICR instruments and resolution limitations in orbitrap instruments. Spectral stitching using spectra obtained with different ion flight times on an FTICR mass spectrometer has been done by Viant and coworkers. This method improved the mass accuracy, sensitivity, and dynamic range of the stitched dataset collected. [12] However, spectral stitching of data from different high resolution mass spectrometers has not been reported.

Here we demonstrate the application of such a stitching method to obtain a comprehensive characterization of lower molecular weight compounds from bio-oil (*m/z* 100-500). The complimentary advantages of high resolving power in FTICR and reduced low-mass discrimination in orbitrap provide a more complete and more accurate molecular characterization of the bio-oil sample.

Methods

Chemicals

Lithium acetate (CAS # 546-89-4) and solvents were purchased from Sigma-Aldrich (St. Louis, MO) at best available purity. The bio-oil samples were provided by Prof. Robert Brown's group at Iowa State University. The bio-oil used for this study was produced by fast pyrolysis of red oak in a pilot-scale reactor located at the Iowa State University Biocentury Research Farm. The sample was recovered from the third fraction of a fractionated bio-oil recovery system. Further details on this system can be found elsewhere.[13] This recovered fraction of bio-oil is dark in color, viscous, and chemically unstable; therefore, to slow chemical transformation, the samples were diluted in methanol at a concentration of 1mg/mL and stored at 4^0C until analysis. Bio-oils were stored in chemically resistant Nalgene bottles. The stock solutions of bio-oil were further diluted to a final concentration of 0.1 $mgmL^{-1}$ in methanol and water (50:50, v/v) prior to analysis.

Mass Spectrometric Data Acquisition

MS analyses were performed using APCI in positive-ion mode. Two different high-resolution mass spectrometers were adopted for the analysis: FTICR (7T Solarix, Bruker, Billerica, MA) and orbitrap (LTQ-Orbitrap Discovery, Thermo Scientific, San Jose, CA). Each instrument was carefully tuned to provide adequate signal and to minimize fragmentation and/or aggregation.

The FTICR was tuned and calibrated with lithium acetate clusters and these parameters were fine-tuned with bio-oil. The ionization parameters used for FTICR were as follows: APCI probe temperature of 350^0C, drying gas temperature of 180^0C with a flow rate of 4.0L/min, nebulizer gas flow rate of 1.0 bar, capillary voltage of 2000V, corona needle voltage of 3000nA, spray shield voltage of -500V and a syringe flow rate of 600 uL/h. The data acquisition size for FTICR was 4 M with a transient length of approximately 1.2 seconds. Two ion flight times (0.4 ms and 0.6 ms) were used to cover the bio-oil mass range of 100-1000 *m/z*.

The orbitrap parameters were tuned with bio-oil. The ionization parameters used were as follows: APCI vaporization temperature of 350^0C, sheath gas flow rate of 50 (arbitrary unit), auxiliary and sweep gas flow rates of 5 (arbitrary unit), capillary voltage of 45V, capillary temperature of 200°C, source voltage of 4650V, source current of 5 µA and tube lens voltage of 80V. The transient length of the orbitrap at the maximum resolution setting of 30,000 is approximately 0.5 seconds.

Data Analysis

The FTICR data was calibrated with known bio-oil compounds in DataAnalysis 4.0 (Bruker, Billerica, MA), converted into a text file, and exported into Composer (Sierra Analytics, Modesto, CA) for further analysis. The orbitrap data was converted into a text file and exported into Composer for calibration and further analysis. The datasets were combined using Excel (Microsoft, Redmond, WA) and the resulting "stitched" dataset was exported to Composer for analysis.

Gel Permeation Chromatography

The HPLC system used for GPC was a Dionex Ultimate 3000 (Sunnyvale, CA) equipped with two Agilent PLgel 3µm 100Å 300 x 7.5mm (p/n PL1110-6320) and one Mesopore 300 x 7.5mm (p/n PL1113-6325) size exclusion columns. A Shodex Refractive Index detector (RI) was used. Bio-oil (0.02g) was dissolved in 10mL tetrahydrofuran and filtered through a Whatman 0.45µ glass microfiber syringe before GPC analysis. Further information about this GPC methodology can be found in previously published work.[14]

Results and Discussion

Individual Spectra

The mass spectrum for bio-oil obtained using the orbitrap (figure 1a) shows ions ranging from *m/z* 100 to 400; however, all of the peaks with relative abundances above 5% are below *m/z* 200. The most abundant ion is found at *m/z* 167.0708 with a chemical composition of $[C_9H_9O_3+H]^+$. The heteroatom class distribution of the

orbitrap data (figure 1b) shows compounds ranging from 1 to 7 oxygen atoms. The most abundant heteroatom class contains 3 oxygen atoms and represents approximately 30% of the overall signal for the identified peaks. The next most abundant heteroatom classes, 2 oxygen atoms and 4 oxygen atoms, are each approximately 15-20% of the overall signal for the identified peaks.

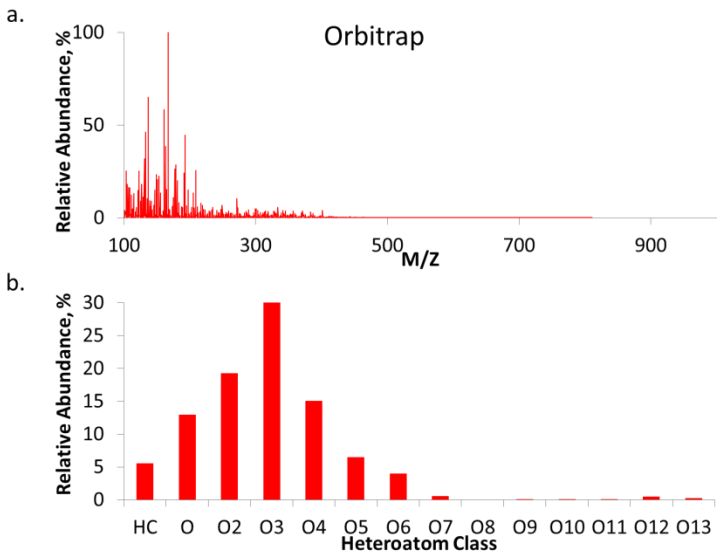


Figure 1. The bio-oil spectrum obtained with APCI-orbitrap (a) along with the heteroatom class distribution created from the APCI-orbitrap data (b).

The FTICR mass spectrum (figure 2a) with an ion flight time of 0.4ms shows an ion distribution that starts at approximately *m/z* 160 and ends at approximately *m/z* 440; however, most of the higher relative abundance peaks fall within the *m/z* range of 220-400. The most abundant ion is found at *m/z* 333.1333 with a chemical composition of $[C_{18}H_{19}O_6+H]^+$. The distribution of heteroatom classes (figure 2b) starts with 1-2 oxygen atoms and extends to 8-9 oxygen atoms. This distribution is centered on the

heteroatom classes that contain 4-6 oxygen atoms; each of these heteroatom classes accounts for about 20% of the overall signal for the identified peaks.

Increasing ion flight time in FTICR from 0.4 to 0.6ms expanded the overall *m/z* range observed in the spectrum (figure 3a). The *m/z* of the mass spectrum starts approximately at *m/z* 200 and the spectrum start to tail off at approximately *m/z* 650. However, the majority of the higher relative abundance peaks still fall within the same *m/z* range of 220-400 as the FTICR lower ion-flight time mass spectrum (figure 2A).

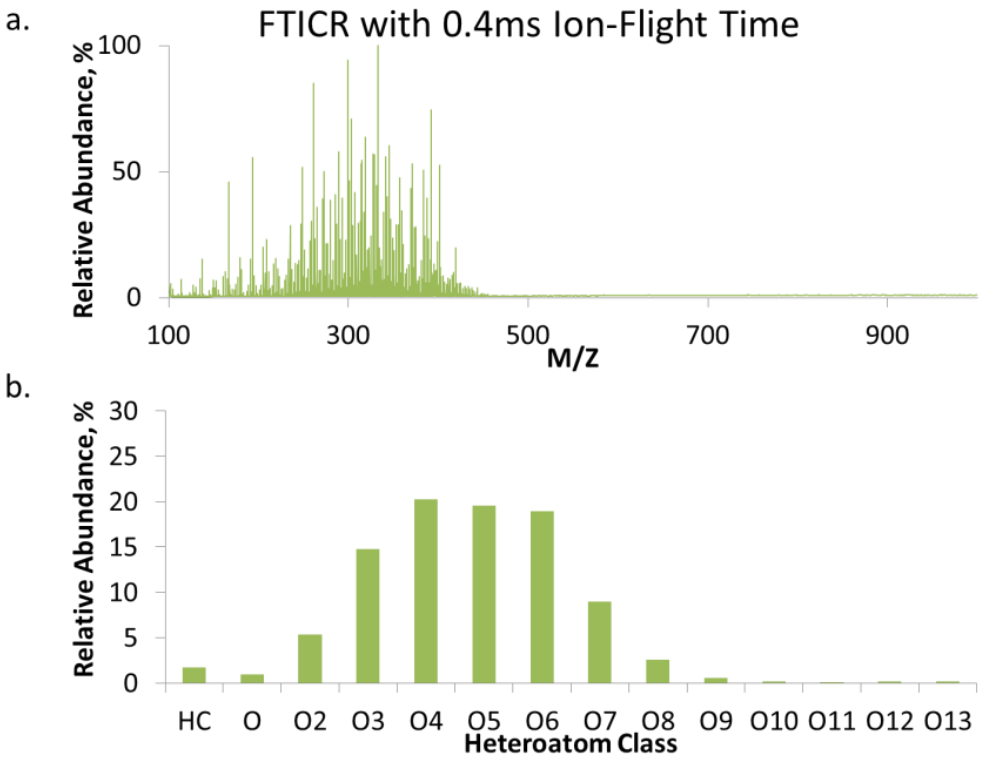


Figure 2. The bio-oil spectrum obtained with APCI-FTICR with an ion-flight time of 0.4ms (a) along with the heteroatom class distribution created from the dataset (b).

The heteroatom class distribution using a longer ion flight time in FTICR shows a wider distribution of classes ranging from 2 oxygen atoms to 12 oxygen atoms (figure 3b).

This wide distribution of oxygen heteroatom classes is centered on the 6 oxygen heteroatom class and accounts for approximately 20% (relative abundance) of the identified peaks.

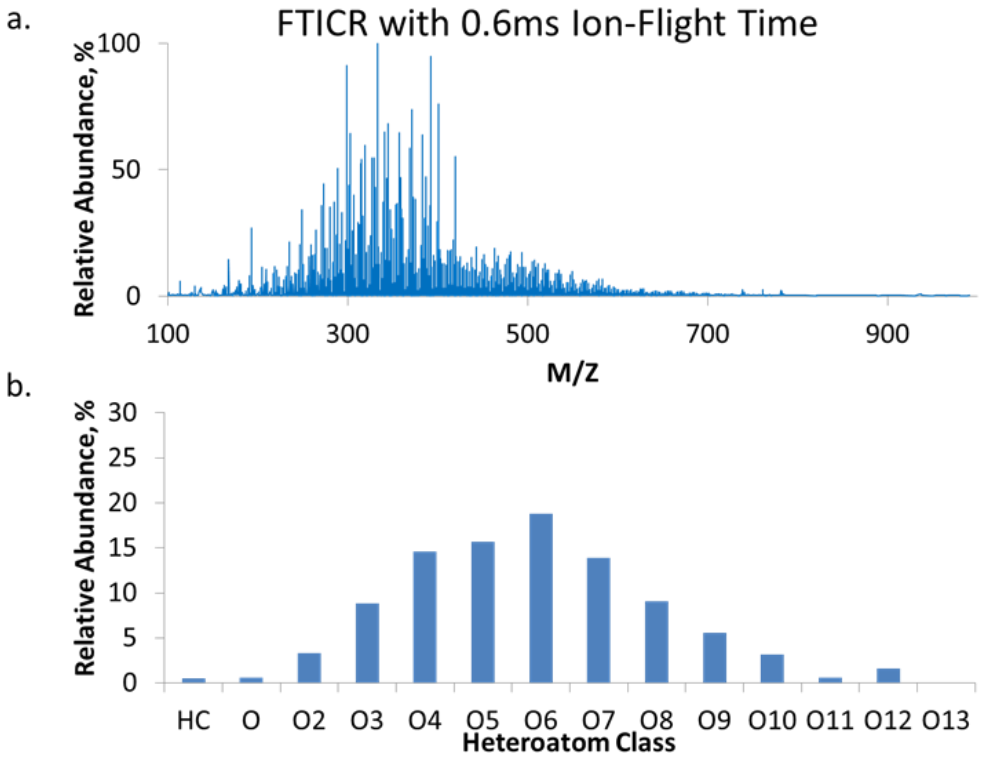


Figure 3. The bio-oil spectrum obtained with APCI-FTICR with an ion flight time of 0.6ms (a) along with the heteroatom class distribution created from the dataset (b).

One important question that needs to be answered is whether high-mass discrimination is present in the FTICR data; specifically high mass discrimination induced by the low ion flight times (0.4ms and 0.6ms) used. Therefore, an Agilent tuning solution was also infused using the same parameters and an ion flight time of 0.6ms to determine whether high-mass ions were excluded due to tuning parameters (figure 4). The Agilent tuning solution has two high-mass ions (*m/z* >400) that were

readily detected with the FTICR instrument parameters. These peaks are observed at *m/z* 622.0287 and *m/z* 922.0102. The detection of these tuning peaks suggests that higher-mass bio-oil compounds, were they present, should be detected (up to at least *m/z* 1000) with an ion flight time of 0.6ms, provided they are efficiently ionized.

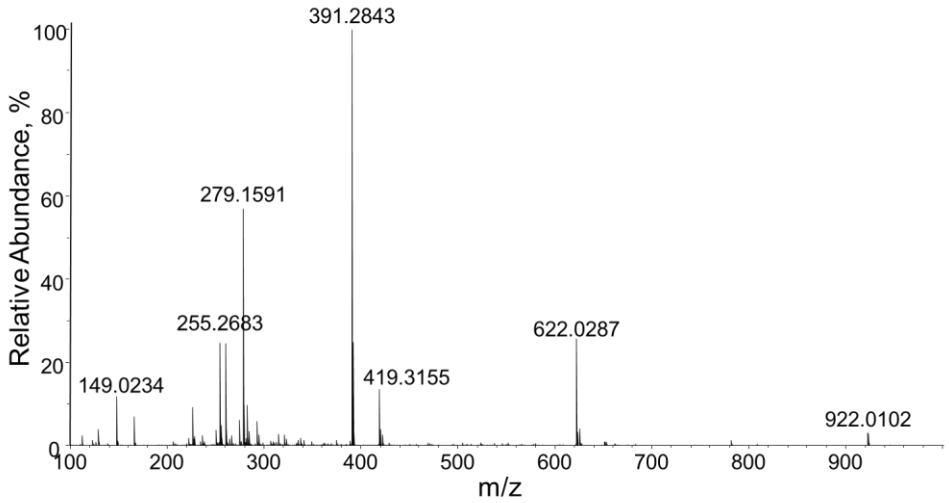


Figure 4. FTICR spectrum of Agilent low-concentration tuning solution collected with an ion flight time of 0.6ms to ensure that ions above *m/z*>400 could be detected.

Stitching Spectra Together

The three MS conditions used for bio-oil analysis yield spectra with a range of ions detected from *m/z* 100-650, with the lower *m/z* range (*m/z*<200) being dominant in the orbitrap data and the higher *m/z* range (*m/z*>200) being dominant in the two FTICR datasets. Therefore, we have stitched together the three datasets to produce a single spectrum that will better represent the molecular makeup of the bio-oil. The spectra were examined to determine peaks that could be used to stitch the datasets together. The peaks used for stitching were selected based on overall relative abundance in each dataset. The selected peaks had close relative abundances in the spectra that were being

stitched together. For example, the peak at *m/z* 193.0880 has a relative abundance of 48% in the orbitrap dataset and a relative abundance of 53% in the 0.4ms FTICR dataset (figure 5). (It is important to note that the relative abundance of these ions is in relation to the most abundant peak in that specific dataset.)

The relative abundances were normalized so that the selected stitching peaks were equal at the specific *m/z* in the spectra that were being stitched together. For example, stitch # 1 in figure 5 normalized the relative abundances in the orbitrap spectrum so that the relative abundance of the peak at *m/z* 193.0770 was equal to the relative abundance in the 0.4ms FTICR spectrum.

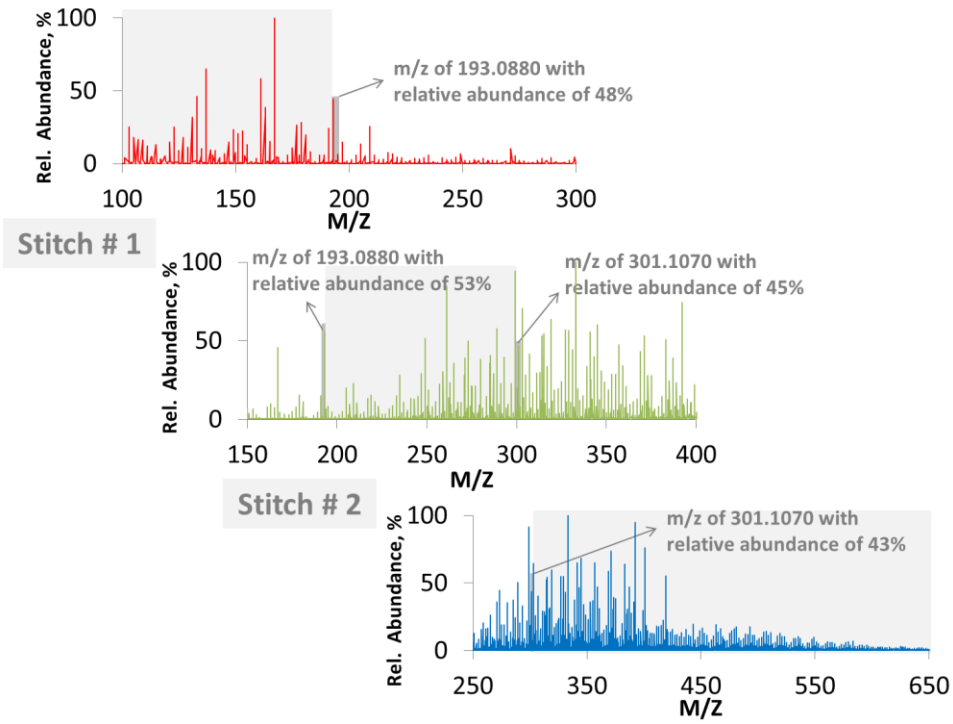


Figure 5. An overview of the spectral stitching methodology. The gray shaded boxes indicate that region of the specific spectrum was used in the final 'stitched spectrum'.

The same procedure was also performed for stitch #2 (figure 5) at *m/z* 301.1070 to stitch the 0.4ms FTICR spectrum with the 0.6ms FTICR spectrum. This results in a stitched spectrum (figure 6) that incorporates the orbitrap data for the ions below *m/z* 193.0770 (shown in red), the 0.4ms FTICR data for the *m/z* range of 193.0770 - 301.1070 (shown in green), and the 0.6ms FTICR data for the *m/z* range > 301.1070 (shown in blue).

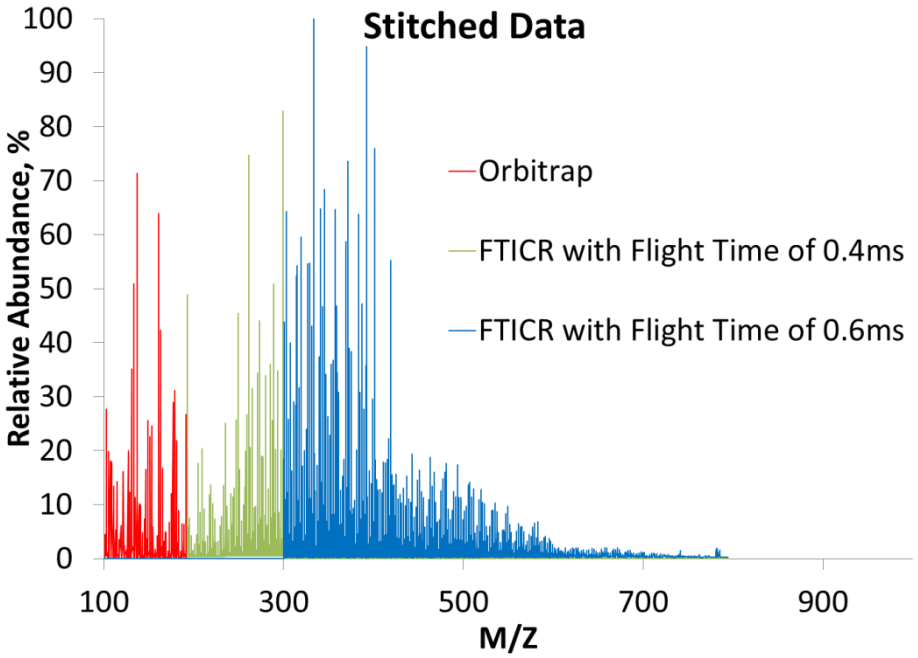


Figure 6. The final "stitched mass spectrum" that is obtained from stitching an orbitrap spectrum with two FTICR spectra obtained at two different ion flight times.

Despite the ability to increase the *m/z* range of bio-oil compounds, there is still a chance that some compounds are absent, under-, or over-represented with this stitching method. Therefore, the bio-oil was analyzed using gel permeation chromatography (GPC) to compare the stitched MS spectrum with the overall size profile obtained by GPC (figure 7). The GPC data shows a molecular weight (MW) range of 100 to 1000Da for this specific bio-oil sample. However, it is important to note that the GPC

MWs were calibrated with a polystyrene standard; and the cross-sectional areas of the calibrants may be quite different from bio-oil compounds, affecting the accuracy of the calibration. For verification, a syringol standard was analyzed, resulting in an observed MW error of 30Da (data not shown). The GPC profile, however, should still be a good global representation of the bio-oil as it should have no mass discrimination; and therefore, a good analytical tool to compare with the stitched spectrum.

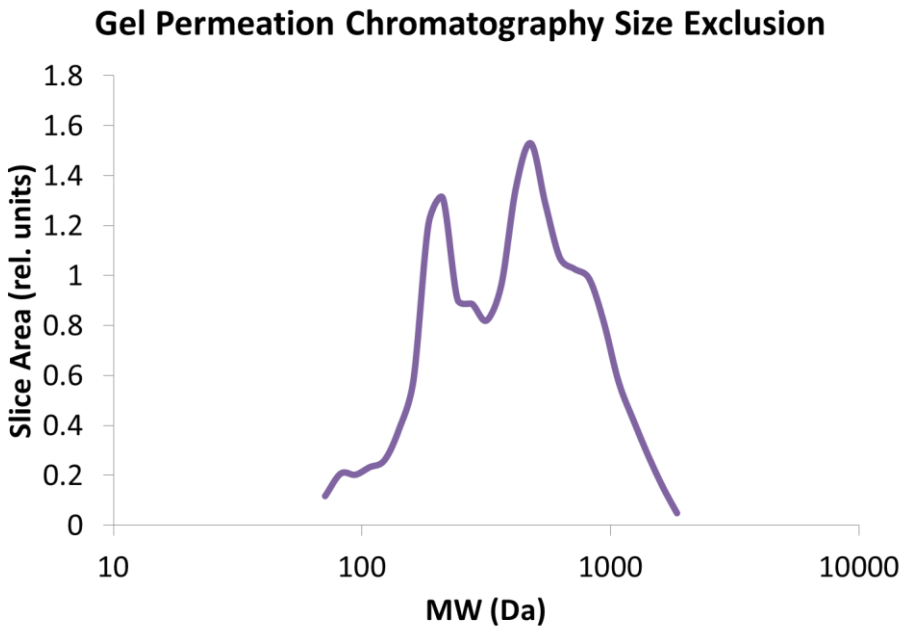


Figure 7. The molecular weight distribution of bio-oil obtained from size exclusion Gel Permeation Chromatography.

The bio-oil GPC profile shows the presence of two MW distributions with local maxima at roughly 200Da and 450Da. The presence of two MW distributions is also shown in the stitched spectrum (figure 6) with local maxima at roughly *m/z* 150 and *m/z* 375. The differences in the local maximum positions between GPC and the stitched mass spectrum are likely due to calibration issues with the GPC. The GPC data shows

that the first MW distribution peak (200 Da) is less intense with a slice area of 1.2 and the second distribution peak (450 Da) with a slice area of 1.4. The slice area ratio between the two MW distributions is 1.2. This is also consistent with the stitched MS data (figure 6) where the relative abundance of the first local maximum (at *m/z* 150) is approximately 65% and the relative abundance of the second local maximum (at *m/z* 375) is approximately 80% and correlates to a relative abundance ratio of 1.2 between the two peaks. This provides evidence that the stitched mass spectrum provides a better profile of bio-oil that is more consistent with GPC data than the unstitched MS spectra would (figures 1-3).

The petroleomic analysis for the stitched bio-oil MS data is shown in figure 8. The heteroatom class distribution (figure 8a) shows a wider range of oxygen heteroatom classes than observed with any individual bio-oil dataset (figures 1-3).

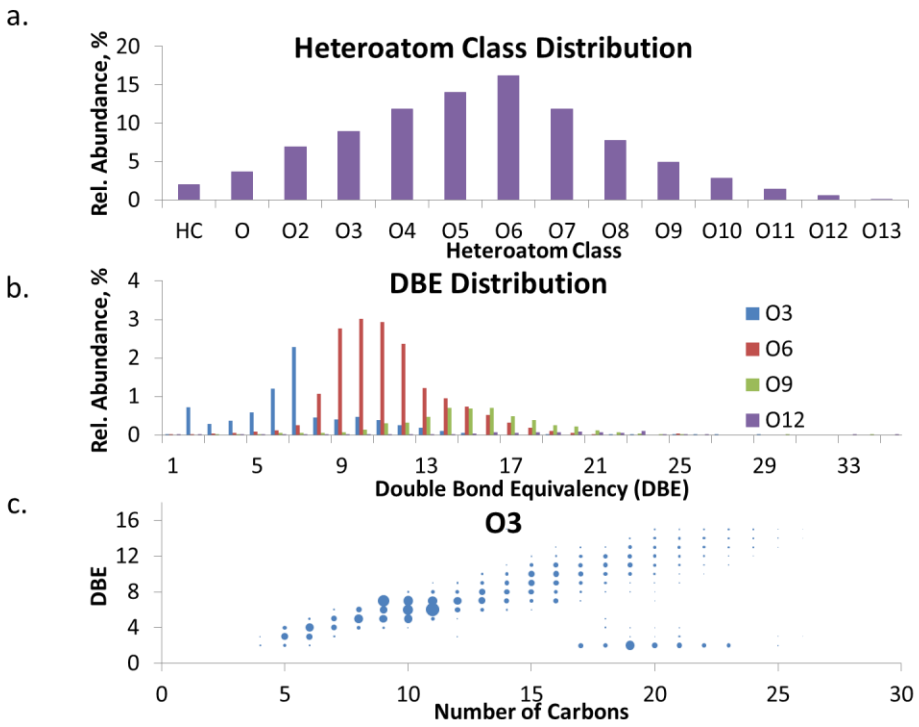


Figure 8. A typical petroleomic analysis of the 'stitched' bio-oil spectrum, showing heteroatom class distribution (a), DBE distribution of a few major heteroatom classes (b), and a contour plot comparing number of carbons vs. DBE values for a specific heteroatom class (c).

The stitched bio-oil data shows oxygen heteroatom classes ranging from 1 oxygen atom to 12 oxygen atoms. The most abundant oxygen heteroatom class contains 6 oxygen atoms and accounts for approximately 15% of the overall signal for the identified peaks. The double bond equivalency (DBE) distribution (b) and a contour plot of carbon number vs. DBE for the 3 oxygen heteroatom class (c) are also shown in figure 8.

Conclusions

A novel data stitching methodology was successful developed to incorporate multiple datasets from two different instrumental platforms. The method was used to overcome mass discrimination in FTICR (low-mass discrimination) and orbitrap (high-mass discrimination). The orbitrap data was used for the low-mass ions (*m/z* <193.0880) and FTICR data at two different ion-flight times, was used for the higher mass ions (an ion-flight time of 0.4ms for 193.0880< *m/z* <301.1070 and an ion-flight time of 0.6ms *m/z* >301.1070). This yielded a single spectrum with ions predominantly in the range of *m/z* 100 to *m/z* 650. The 'stitched spectrum' was also compared with GPC and this showed that the stitched spectrum was a better representation of the bio-oil compared to the individual datasets obtained from orbitrap and FTICR.

References

1. T. Marker, *Opportunities for biorenewables in oil refineries*, UOP LLC, 2005.

2. M. S. Mettler, D. G. Vlachos and P. J. Dauenhauer, *Energy & Environmental Science*, 2012, **5**, 7797-7809.

3. E. A. Smith and Y. J. Lee, *Energy & Fuels*, 2010, **24**, 5190-5198.

4. E. A. Smith, S. Park, A. T. Klein and Y. J. Lee, *Energy & Fuels*, 2012, **26**, 3796-3802.

5. D. Paul Cole, E. A. Smith, D. Dalluge, D. M. Wilson, E. A. Heaton, R. C. Brown and Y. Jin Lee, *Fuel*, 2013.

6. A. G. Marshall and R. P. Rodgers, *Acc. Chem. Res.*, 2004, **37**, 53.

7. M. P. Barrow, L. A. McDonnell, X. Feng, J. Walker and P. J. Derrick, *Analytical Chemistry*, 2003, **75**, 860-866.

8. K. O. Zhurov, A. N. Kozhinov and Y. O. Tsybin, *Energy & Fuels*, 2013, **27**, 2974-2983.

9. A. G. Marshall and C. L. Hendrickson, *Annu. Rev. Anal. Chem.*, 2008, **1**, 579-599.

10. N. K. Kaiser, G. E. Skulason, C. R. Weisbrod, S. Wu, K. Zhang, D. C. Prior, M. A. Buschbach, G. A. Anderson and J. E. Bruce, *Rapid communications in mass spectrometry*, 2008, **22**, 1955-1964.

11. M. Hardman and A. A. Makarov, *Analytical Chemistry*, 2003, **75**, 1699-1705.

12. A. D. Southam, T. G. Payne, H. J. Cooper, T. N. Arvanitis and M. R. Viant, *Analytical Chemistry*, 2007, **79**, 4595-4602.

13. A. Pollard, M. Rover and R. Brown, *Journal of Analytical and Applied Pyrolysis*, 2012, **93**, 129-138.

14. P. A. J. Marjorie R. Rover, Ryan G. Smith, Robert C. Brown, *Submitted to Bioresource Technology*, 2013.

CHAPTER V

REAL-TIME MONITORING OF BIOMASS FAST PYROLYSIS WITH MICRO-PYROLYSIS-ATMOSPHERIC PRESSURE CHEMICAL IONIZATION TIME-OF-FLIGHT MASS SPECTROMETER

A manuscript to be submitted to *Energy and Environmental Science*

Erica A. Smith, Carolyn Hutchinson, D. Paul Cole, and Young-Jin Lee*
Department of Chemistry, Iowa State University, Ames, IA 50011

Broader Context

Biomass fast pyrolysis is a promising and sustainable technique that thermochemically converts lignocellulosic biomass into vapors. These vapors can be condensed into an energy-dense liquid or bio-oil and bio-oil can be upgraded into biofuels and/or other chemicals to supplement and/or replaced petroleum-based fuels and products. However, yields of the condensable vapors vary greatly depending on the pyrolysis operating conditions and biomass make-up. Additionally, very little is known about the pyrolysis process itself and the mechanisms and kinetics involved, specifically at the molecular-level. Therefore, we describe a new analytical technique that allows for the molecular constituents of pyrolysis to be monitor in real-time. This technique opens up opportunities for fundamental studies that have been previously difficult or impossible to do. The biggest impact of this technique will be on fundamental and non-global mechanism studies that would benefit from being able to identify and monitor chemical species during pyrolysis. And having a better grasp of the individual mechanisms involved in pyrolysis will led to a better understanding of the

kinetics which is invaluable information for better controlling pyrolysis and it's end-products.

Abstract

A fundamental understanding of biomass fast pyrolysis is vital in order to optimize the thermochemical conversion process and allow for it to be a viable renewable energy platform. More specifically, the lack of molecular-level information on fast pyrolysis limits the ability to construct appropriate mechanisms to use for determining the kinetics of fast pyrolysis. Here we present a novel technique that can monitor fast pyrolysis products at the molecular-level in real-time. A fast-scanning time-of-flight mass spectrometer is coupled with a soft ionization technique and a drop-in microfurnance with a well-defined start time of pyrolysis. This provides insight into biomass fast pyrolysis that is not possible with traditional techniques. For example, metastable intermediates of cellulose pyrolysis could be identified and monitored with this novel approach. This platform also showed that furan formation actually competes with levoglucosan formation in cellulose pyrolysis based on furans' extract ion chromatograms. Fundamental studies are also possible including studying the differences in pyrolyzate time profiles between different attributes of single red oak particles (rod-shape vs. flake shape).

* To whom correspondence should be addressed. E-mail: yjlee@iastate.edu

Introduction

Biomass pyrolysis, specifically fast pyrolysis, is a promising technique for renewable energy for transportation fuels and other petroleomic-based chemicals.[1] Many different biomass materials can be used for this process; this includes any organic matter and encompasses woody and herbaceous species, plant waste, agricultural waste and industrial waste.[2] Most biomass is a composite of three biopolymers (cellulose, hemicellulose, and lignin) and their extractives with each respective amounts of biopolymer being unique to the specific biomass being utilized. For example, most woody biomasses, like red oak, contain roughly 40-50% cellulose, 25%-35% hemicellulose, and 20-40% lignin.[3]

Pyrolysis is a thermochemical conversion of biomass that operates at moderate temperatures (usually between 400-600°C) and with little or no oxygen present. Fast pyrolysis involves rapidly heating biomass (> 1200°C s^{-1}) at roughly atmospheric pressure and with a short residence time (1-5 s). [4] The products of pyrolysis are chars, non-condensable gases, and condensable vapors; the condensable vapors are collected as bio-oil, which can be further upgraded into drop-in fuels.[4] Bio-oil can be characterized by anhydrosugars, phenolics, furans, and small oxygenates (like acetic acid and formic acid). These compounds also include a wide range of non-volatile and volatile species.[5]

Fast pyrolysis has been shown to produce various amounts of bio-oil depending on reaction conditions, [3] and a high heating rate and short residence time are critical factors in the formation of condensable vapors. The formation of condensable vapors is kinetically favored with a high heating rate and short residence times. In comparison,

slow pyrolysis operates at a lower temperature and longer residence time where char and non-condensable gases are the major products; this suggests that the formation of char and non-condensable gases are thermodynamically favored.[6] It is thought that a better understanding of the pyrolysis process would allow for better control of the reactions. Better control of the reactions could help decrease the complexity of the mixture or allow for more desirable compounds and/or make the bio-oil more valuable.

Studying the kinetics and the mechanisms of biomass pyrolysis is important to be able to control the end products. This is why much attention has been spent on studying pyrolysis kinetics. Currently, most of the kinetic measurements have been made by thermogravimetric analysis (TGA) and micropyrolysis-gas chromatography-mass spectrometry (μPy-GC-MS).[7-10] TGA has focused on obtaining more of a global or semi-global mechanism for specific biopolymers of biomass. This technique suffers from low heating rates that do not accurately resemble fast pyrolysis. At lower heating rates each biopolymer (cellulose, hemicellulose, and lignin) has a unique decomposition temperature range.[11] However, at higher heating rates, like in fast pyrolysis, the decomposition of the biopolymers is less distinguishable making TGA a less effective tool for studying the kinetics and mechanisms of fast pyrolysis. And the use of a GC separation in traditional μPy-GC-MS eliminates the possibility of time-resolved pyrolysis data for the identified compounds. Also, some of the pyrolyzates may be metastable compounds that may not survive in the GC column.

There has been much ambiguity in the reported kinetic data obtained using global mechanisms. For example, experiments done on cellulose, the most simple and

abundant biopolymer of biomass, has shown conflicting results in the experimental kinetic parameters.[8] Most of these experiments cannot distinguish between vastly different chemical compounds; and therefore, they cannot explain the individual degradation mechanisms that can affect the overall kinetics.

Research has been conducted to address the lack of non-global mechanisms and low heating-rate experiments. Lédé and coworkers developed an innovative technique that utilizes a focusing 5kW xenon lamp to conduct radiant flash pyrolysis.[12] The xenon lamp allows for a controlled heating time (down to 10 ms) and for rapid quenching of gas and liquid intermediates. The quenched intermediates are then collected and analyzed ex-situ using GC-MS and LC-MS. However, one limitation of this technique is that the exact pyrolysis temperature is unknown. In addition, the analysis is limited to an ex-situ analysis of the pyrolyzates. Hence, it cannot explain time-dependent changes of the pyrolyzates that are produced at a certain time frame and escape from the reactor.

Another novel technique adapted by Dauenhauer and coworkers utilizes high-speed photography to capture images of cellulose pyrolysis.[13, 14] The high-speed photography allowed for the dynamic nature of fast pyrolysis and confirmed the presence of a liquid intermediate also known as molten biomass. A liquid intermediate, or active cellulose, has also been previously suggested by the Broido-Shafizadeh model.[15] The presence of pyrolyzates being ejected as aerosols was also confirmed with high-speed photography.[14] Silylation of the collected aerosols followed by GC-MS analysis indicated that they are mostly cellobiosan dimers formed through direct decomposition of cellulose and transported via aerosol ejection. Dauenhauer and

coworkers also developed a thin-film (cellulose) pyrolysis technique that eliminated the conduction effects in pyrolysis.[8] The thin-film allowed for isothermal kinetics to be explored and showed that α-cyclodextrin could serve as a surrogate compound for modeling cellulose pyrolysis.

Fundamental studies, like Dauenhauer and Lédé, have revealed the pyrolysis process to be much more complex than can be explained by global mechanisms or simple kinetic models. The presence of a liquid intermediate and aerosol ejection suggests that pyrolysis is not simply controlled by kinetics but also influenced by the dynamics of the system. In order to construct a more encompassing mechanism, there needs to be more fundamental studies that uncover the molecular details of the vapor phase intermediates.[16] However, characterization of the vapor phase intermediates has largely been absent due to the lack of analytical tools that can study the intermediate species.

The vapor phase pyrolysis products are eventually condensed and collected as bio-oil. Bio-oil is a complex mixture of non-volatile, volatile, polar, and nonpolar compounds.[5] Our group has done extensive and pioneering research in developing high resolution mass spectrometry (HRMS) techniques to perform bio-oil analysis.[17-19] HRMS provides accurate mass measurements that allow for confident assignments of chemical compositions and soft ionization techniques have been adapted for the analysis of non-volatile and/or thermally unstable compounds. Because of these strengths, most of our efforts were focused on analyzing compounds that are not amenable to analysis with traditional techniques, like GC-MS. For example, hundreds of non-volatile lignin

dimer and trimer compounds were characterized with laser-desorption ionization (LDI) HRMS.[17] With negative mode electrospray ionization (ESI) and HRMS, we could characterize over eight hundred volatile and non-volatile bio-oil compounds.[18] The accurate chemical composition assignments reveal the double-bond equivalent (DBE); therefore, phenolic compounds (DBE $\geq$ 4) can be distinguished from holocellulose-derived aliphatic compounds (DBE < 4). This technique has been applied for the speciation of nitrogen-containing compounds in switchgrass bio-oil and determined the structural motif of these nitrogen compounds.[19]

Here we expand the capability of HRMS by connecting a micropyrolyzer to the HRMS and monitor the vapor phase of pyrolysis products in real-time. We adopt a "drop-tube" micropyrolyzer, which has a high heating rate (up to 1000°C/s), and atmospheric pressure chemical ionization (APCI), which ionizes non-volatile molecules with no or minimal fragmentation. We applied this novel instrumentation to the pyrolysis of cellulose and red oak biomass, and traced the changes in pyrolysis products with high temporal resolution. This study unveils fine details of fast pyrolysis that was previously unattainable.

Methods

Materials and Sample Preparation

Levoglucosan and other model bio-oil compounds (hydroxymethyl furfural, ribose, and glucose) were purchased from Sigma-Aldrich (St Louis, MO) at their highest available purity (at least 98%). Sigmacell Cellulose Type 20 (20 μm) was also

purchased from Sigma-Aldrich for the cellulose powder and thin-film pyrolysis experiments. Northern red oak was obtained from Wood Residuals Solutions (Montello, WI); the red oak was ground and sieved for a size less than 200 µm by our collaborators from the Center of Sustainable Environmental Technologies at Iowa State University.

Due to the extremely high sensitivity of the APCI-Time-of-Flight (TOF) mass spectrometer, the amount of cellulose powder and red oak being pyrolyzed had to be very small to prevent the saturation of the TOF detector. (Typically, we loaded approximately 16 µg.) For the red oak experiments, only one particle of red oak (typically 50 µg) was added to the pyrolysis cup for analysis.

A thin-film of cellulose was prepared using the procedure described by Dauenhauer and co-workers.[20] Namely, cellulose (Sigmacell 20µm, Sigma Aldrich, St. Louis, MO) was suspended in water at a concentration of 1 mg mL^{-1} and 25 µL of cellulose suspension was quickly pipetted into a deactivated stainless steel pyrolysis cup (4 mm x 8mm). The water was removed by evaporation at 40 °C in an isotemp oven. The thin-film thickness is expected to be around 3 µm according to the calculation and measurement by Dauenhauer and co-workers with SEM.[20]

Experimental Design

A single-shot drop-in micro-furnace pyrolyzer (Frontier Laboratories 2020iS Micropyrolyzer, Japan) set at 500°C was installed onto an Agilent 7890A GC oven (see Figure 1 for instrument set-up). Installed in the GC oven was 0.60 meters of deactivated fused silica (i.d. 100 µm and o.d. 360 µm; SGE Analytical Science, Austin, TX) between the injection port and APCI interface; this was used to transport the pyrolyzates

from the micropyrolyzer (µPy) to the mass spectrometer with minimal dead time and volume. The dead time for the deactivated fused silica is calculated to be 0.5 seconds with the flow rate (1 ml min-1) used in this experiment. The temperature of the GC inlet, oven, and interface were all set to 320 °C. Ultra high purity helium gas was used as a carrier gas with a flow rate of 100 mL/min through the pyrolyzer; however, the gas flow was split 100:1 at the GC inlet for a resulting flow rate of 1 mL/min through the capillary.

The GC was coupled to an Agilent 6200 TOF MS with an APCI interface. High purity ammonia gas (Praxair, Des Moines, IA, USA; 500 ppm in helium) was introduced into the APCI chamber at a flow rate of 1 mL/min through a tee that was installed inside the GC oven right before the APCI interface. The ammonia gas is flowing outside the deactivated fused silica and comes in contact with the pyrolyzates at the exit of the capillary within the APCI ion source chamber. The auxiliary heater on the APCI interface was held at 320°C to prevent the pyrolyzates from condensing out before reaching the APCI chamber.

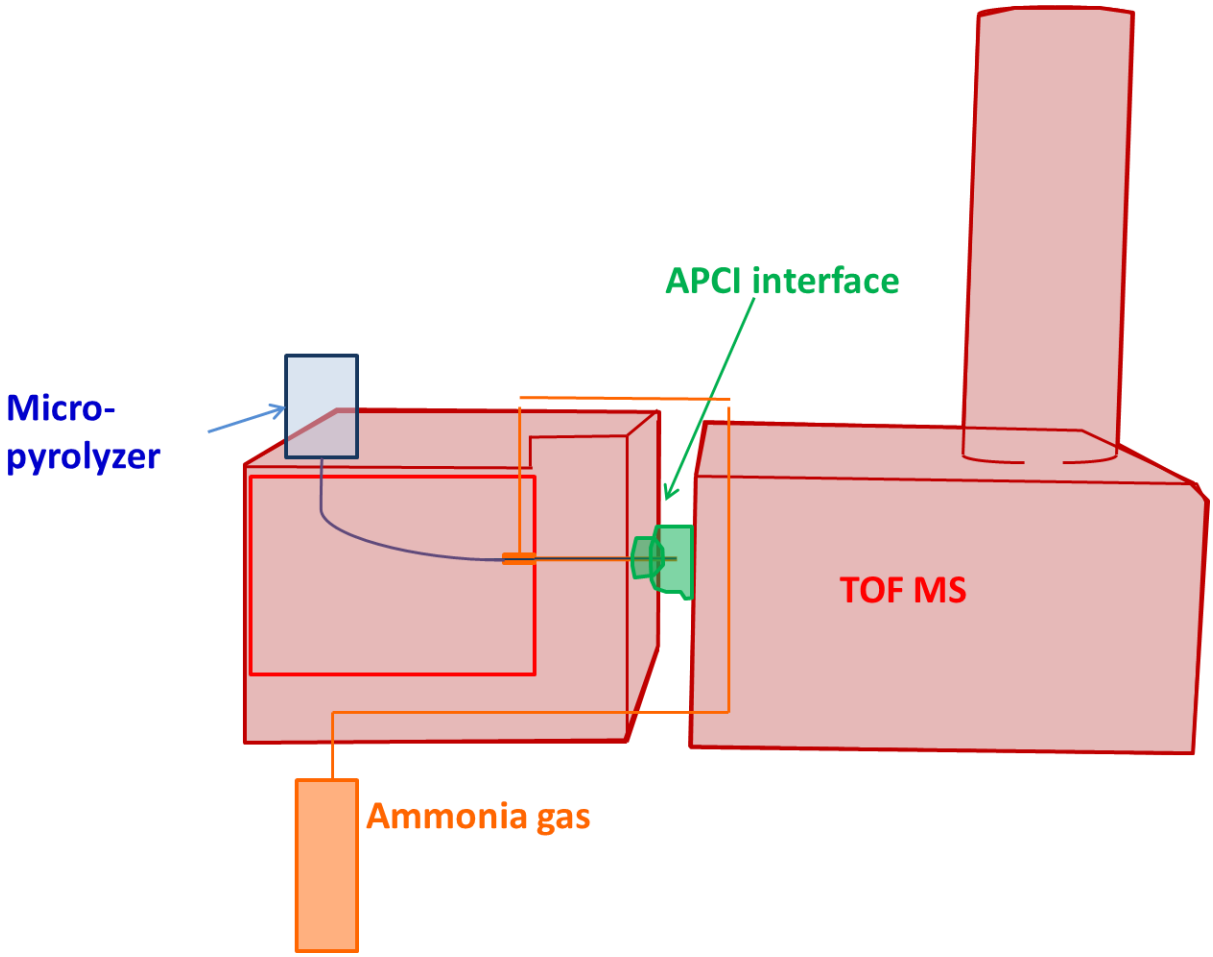


Figure 1. A schematic of the instrument set-up for the µPy-APCI-TOF.

Mass Spectrometer Parameters

Mass spectrometer parameters were optimized with bio-oil model compounds (a mixture of phenols, sugars and furans). The drying gas flow rate was set to 5 L/min with a temperature of 350°C. The corona was set to 1 µA and the voltage on the capillary cap (VCap) was set to 1000 V. This produced a corona discharge voltage of approximately 4200 V. The fragmentor and skimmer voltages were optimized to ensure that fragmentation and/or aggregation of the compounds would not occur. The optimal voltages for the fragmentor and skimmer were found to be 95V and 65V, respectively.

The TOF MS was setup to scan an *m/z* range of 100-1000 with a scan rate of 24 spectra per second.

Results and Discussion

The novel analytical platform, µPy-APCI-TOF, we developed (figure 1) is significantly different from the traditional analytical techniques, like µPy-GC-MS, and allows for the monitoring of pyrolysis at the molecular level and in real-time. First of all, we bypass the GC separation and directly connect the pyrolyzer to the mass spectrometer with a minimal dead time (approximately 0.5 seconds) in the transfer line. Instead of using a GC separation we reply on high-resolution MS separation with a soft ionization of the analyte molecules. This approach has several advantages; 1) fast analysis of the pyrolyzates providing their time-depended changes, 2) ionization of non-volatile molecules not amendable to typical GC-MS analysis (e.g. lignin oligomers), and 3) direct chemical composition analysis of each pyrolyzates. We adapted a time-of-flight mass spectrometer (TOF-MS) as a HRMS system because of its fast scanning speed (1 mass spectrum ≤ 1 ms, we used 24Hz in this study). Additional, a well-define time zero of pyrolysis is an essential element for this study and a drop-in microfurnace is well suited for this purpose. We first tested and optimized the system with a simple mixture of model compounds, then applied to cellulose and to single particles of biomass.

<u>Dopant-assisted APCI for Model compounds without the pyrolyzer</u>

An initial experiment was conducted on the APCI-TOF without the pyrolyzer by injecting levoglucosan into the GC injection port (figure 2a). The primary focus of this experiment was to optimize MS experimental conditions with a known compound to ensure that pyrolyzates would not fragment or aggregate at the MS interface. When the experiment was performed without a dopant gas the protonated ion of levoglucosan at *m/z* 163.061 (Δm_{error} = 3 ppm) is observed as well as fragment ions at *m/z* 145.050 and *m/z* 127.039. The most abundant ion, at *m/z* 145.050, corresponds to the water loss of levoglucosan and the next most abundant fragment ion, *m/z* 127.039, corresponds to a double water loss of levoglucosan. The significant presence of fragment ions suggests that ambient APCI provides sufficient internal energy to the protonated levoglucosan to cause fragmentation.

Dopant-assisted APCI has been shown to aid the ionization of various compounds and the use of ammonia gas as a dopant has been shown to prevent fragmentation of analytes in LC-MS analysis.[21] Ammonia gas is present at a high concentration at the APCI interface and preferentially ionized, producing ammonium ions through self-reaction. These ions then produce proton or ammonium ion adducts with the analytes and has minimal internal energy transfer with little fragmentation. When ammonia gas was introduced to the APCI chamber, ammonium-adducted levoglucosan was observed at m/z 180.087 (Δm_{error}= 1 ppm, figure 2b). The addition of ammonia gas led to a 20x increase in signal for the ammonium-adduct levoglucosan compared to the protonated levoglucosan without ammonia dopant gas.

Ion signals for levoglucosan show a linear response ($R^2 = 0.9868$) for a concentration range of 0.6 µmol to 15 mmol (Figure 2c); thus, demonstrating the wide dynamic range of this instrument. The limited of detection is determined to be approximately 60-1000 nM, which is a much lower limit of detection compared to the reported value for small metabolites in GC-APCI-TOF.[22] An equimolar mixture of bio-oil model compounds (0.1mM each) was injected through the GC injector into the system to determine their ionization efficiencies compared to levoglucosan. The compounds that represent cellulose/hemicellulose pyrolysis products (levoglucosan, glucose, hydroxymethyl furfural (HMF), and ribose) showed only ammonium-adducted ions; whereas, the lignin pyrolysis model compounds (acetosyringone, syringaldehyde, and vanillic acid) show both protonated and ammonium adducts (Supplementary Figure 2). Overall, the response factors for HMF, glucose, ribose, and vanillic acid are much lower compared to levoglucosan, syringaldehyde, and acetosyringone.

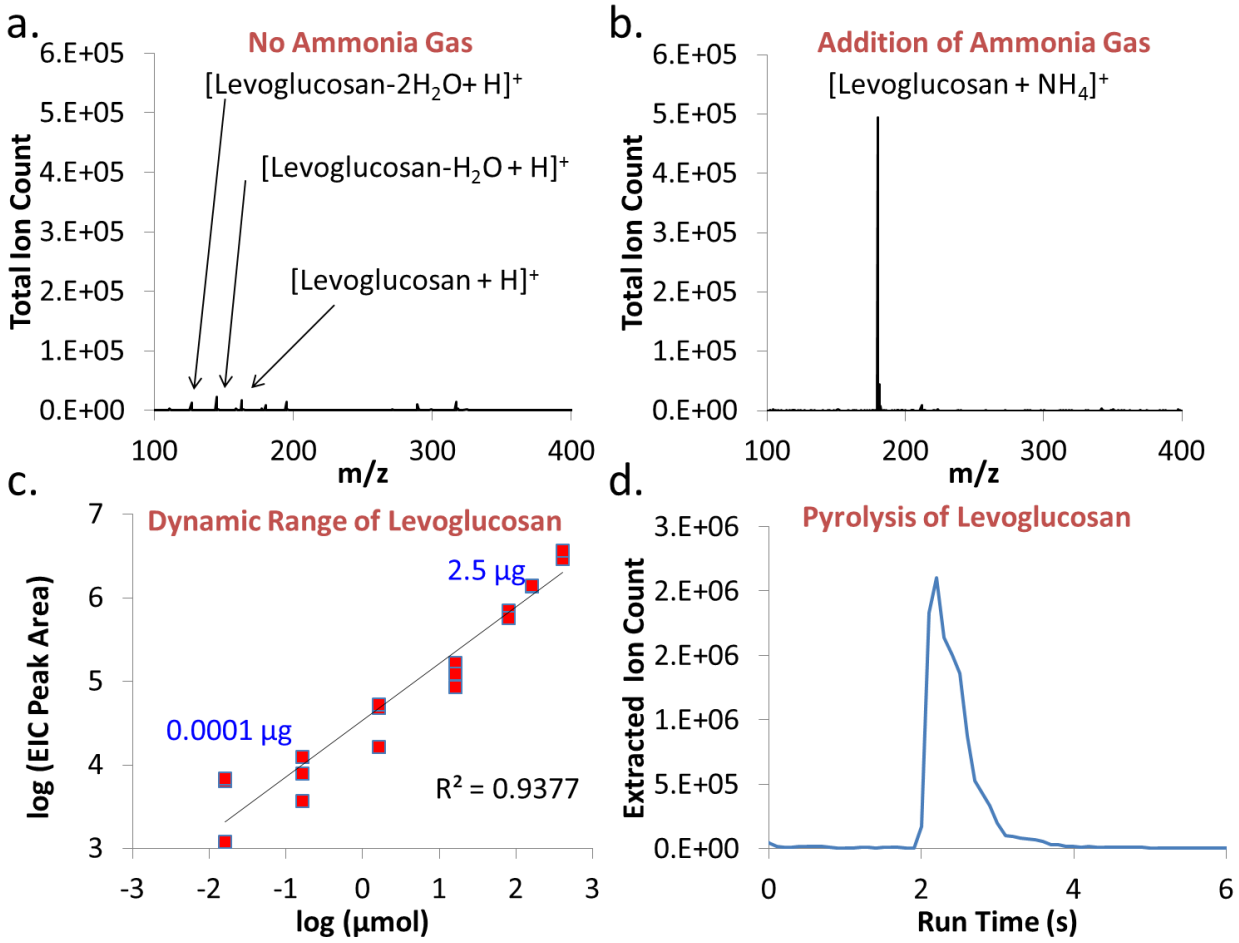


Figure 2. The optimization of the APCI-TOF was done using levoglucosan without the pyrolyzer (a-c) and with the pyrolyzer (d). These basic experiments showed that the use of ammonia gas as a dopant gas (b) minimizes the in-situ fragmentation that is observed without ammonia dopant gas (a) and dramatically increases the sensitivity and dynamic range (c). The pyrolysis of levoglucosan also shows the time-resolution capabilities that are possible with the system (d).

After we had successful developed and tested the dopant-assisted APCI for the soft ionization of model pyrolysis compounds, we installed the drop-in microfurnace pyrolyzer and tested the full system with levoglucosan. We dropped a stainless steel cup with approximately 10 µg of levoglucosan into a 500°C microfurnance as we were

acquiring TOF-MS data. The levoglucosan peak starts to appear around 1.8 seconds after the cup is dropped and has a band broadening of 0.5 seconds (figure 2d). According to our calculation, the drop time is ~0.2 seconds and the transit time is ~ 0.5 seconds; hence, the rest of the ~1.0 seconds is considered to be data acquisition and/or dead time in the mass spectrometer and data delivery time to the computer. The band broadening of 0.5 seconds is much broader than the calculated thermal induced diffusion broadening (~0.004s for levoglucosan at 320°C to travel 0.6 meter of transfer line) and is attributed to several imperfections in the current system, such as carrier gas flow turbulence and injector dead volume. It should also be noted that there is currently ~0.2-0.3 seconds of uncertainty arising from the lack of synchronization between manual dropping and MS data acquisition.

Cellulose Pyrolysis

The first system we studied was the pyrolysis of cellulose powder (figure 3). A deactivated stainless steel cup with approximately 16 micrograms of cellulose powder is dropped into a microfurnance and the pyrolysis products are monitored in real-time with APCI-ToF. Considering the dead time observed in loading levoglucosan into the same system (2d) and the appearance of known cup contaminants, we calculate the actual heating/pyrolysis time to average around 0.5 seconds ± 0.3. The dead time of the system estimated from the loading of levoglucosan also matches with the elution of known pyrolysis cup contaminations. Namely, a minimal amount of volatile cup contaminations are immediately heated and transferred to the MS prior to pyrolysis and can be used for a dead time calibration. The approximate dead time/data acquisition

time can also be monitored with the elution of known pyrolysis cup contaminations. For the remainder of this paper, the time scale in the extracted ion chromatograms (EIC) was calibrated with the time that the pyrolysis cup contaminants first appear.

The total ion count (TIC) for a cellulose pyrolysis run is shown in Figure 3a. Summed spectra for the time first 10 seconds of pyrolysis is shown in Figure 3b. The most abundant peak is found at *m/z* 180.086 which corresponds to the ammonium adduct of $C_6H_{10}O_5$ (Δm= 6 ppm), this is most likely levoglucosan (1,6-anhydroglucopyranose) or its structural isomer (1,6-anhydroglucofuranose). Because of the absence of a chromatographic separation, we cannot distinguish structural isomers; however, we can directly assign chemical compositions from accurate mass information. Also, in cellulose pyrolysis levoglucosan is known to have a much higher yield (15x) than 1,6-anyhdroglucofuranose [8] and thus most of the *m/z* 180.086 peak is expected to come from levoglucosan.

One peak that was quite interesting is found at *m/z* 240.110 with a chemical formula of $[C_8H_{14}O_7+NH_4]^+$ (Δm= 7 ppm). This specific compound, at *m/z* 240.100, has not been observed in traditional µPy-GC-MS experiments or in the typical analysis of cellulose pyrolysis oil. This peak was not found in a µPy-GC-APCI-TOF experiment in the same system but where the deactivated fused capillary tubing was replaced with a J&W DB-1701 column (Supplementary Figure 3). This indicates that the specific compound, $C_8H_{14}O_7$, is a metastable compound with a short life-time as will be discussed later in this section.

Several other peaks, not identified with GC-MS, were also observed with this technique. These are seen at *m/z* 222.097 and 324.129 with chemical compositions of $[C_8H_{12}O_6+NH_4]^+$ and $[C_{12}H_{18}O_9+NH_4]^+$, respectively. Several other compounds are also observed with much lower ion signals, especially below m/z 150 (inset of Figure 3b). The most abundant peak seen in the m/z range of 100 to 150 is at m/z 144.066 which could correspond to levoglucosenone, maltol, and/or hydroxymethyl furfural. Table 1 is a list of all the major m/z ions with their tentative assignments. A total of 50 chemical compositions were identified as products of cellulose fast pyrolysis, which is much greater than the approximately 20 chemical compositions that are identified in typical µPy-GC-MS of cellulose. The pyrolysis compounds or pyrolyzates in Table 1 can be largely grouped based on their carbon ring structure: furan-based (4 carbon ring), levoglucosan- based (5 carbon ring), cellobiosan-based (dimeric; cellobiosan is a 5 carbon ring with a 6 carbon ring) and intermediates between levoglucosan and cellobiosan. Oligomers greater than cellobiosan have been reported in the ESI analysis in cellulose cotton paper; [23] however, oligomers were not observed in this study.

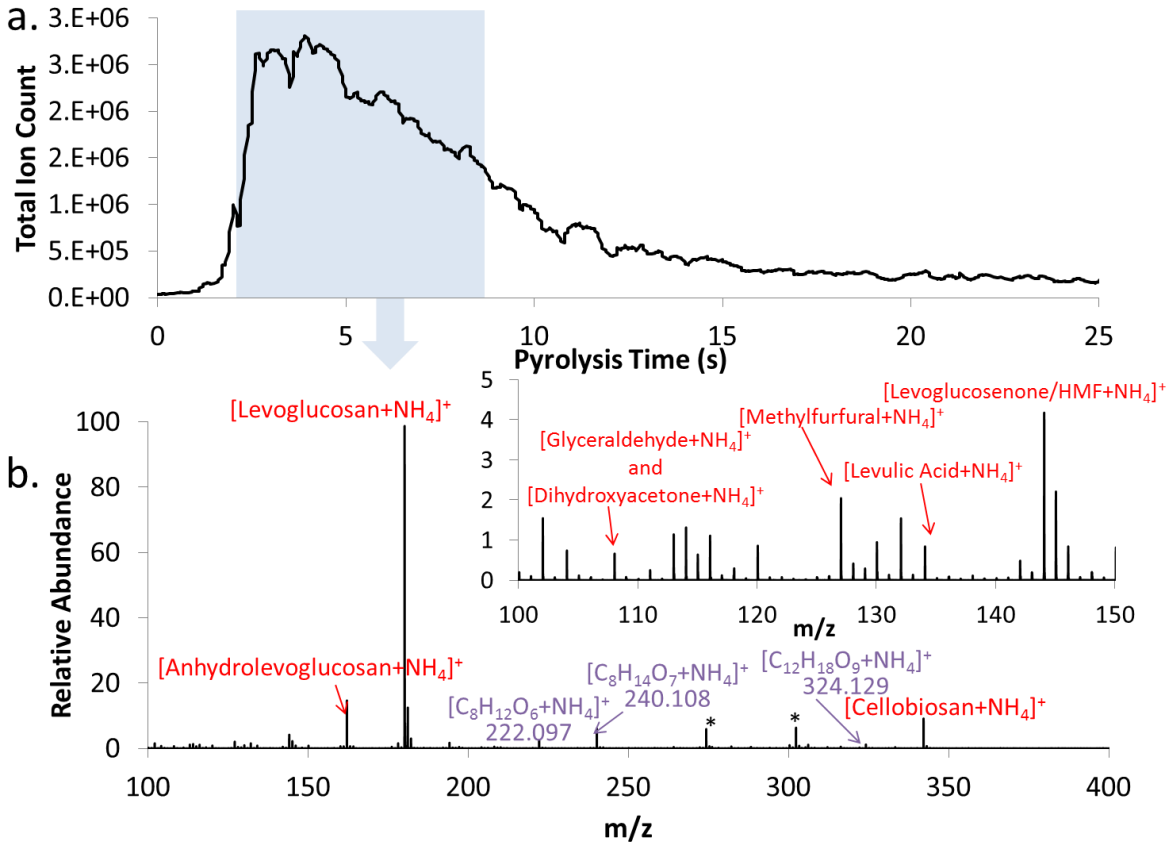


Figure 3. The time profile of the total ion count in real-time monitoring of cellulose pyrolysis (a) and the summed mass spectrum (b) obtained from the region where most of the cellulose pyrolyzates eluted. Also, a zoomed in insert for m/z 100-150 is shown in (b). Some peaks were labelled with their most abundant structural isomer as seen in literature. The (*) denotes pyrolysis cup contaminates that were observed. Time zero is calibrated to represent pyrolysis time as designated from the appearance of volatile contaminates present on the cup..

Further insight about cellulose pyrolysis can be gained when the time profiles of pyrolyzates are extracted via extracted ion chromatograms (EICs). This allows for a specific *m/z* to be monitored during the very short pyrolysis time (< 1 sec) and is a unique capability of the instrument set-up. Various EICs for cellulose pyrolysis are showed in figure 4 and are compared to the total ion count (figure 4a) The time profile of

levoglucosan (figure 4b) shows two distinct profiles, one peak at around 2 seconds and the other peak at 4-5 seconds. The two distinct profiles could be caused by two unique kinetic pathways for levoglucosan formation and/or structural isomers that have a different kinetic pathway than levoglucosan.

Table 1. List of all major m/z ions in cellulose pyrolysis and their tentative chemical composition assignments. The tentative assignments are based on compounds observed in µPy-GC-MS.

m/z	Chemical Composition	Relative Abundance (% of Levoglucosan)	Tentative Chemical Compositions Assignments
100.076	C_5H_6O	0.2	a-methyl furan
102.056	$C_4H_4O_2$	1.6	2-(5H) Furanone
102.092	C_5H_8O	0.1	cyclopentanone, 4-methyl-2,3-dihydrofuran, 2-butenal, 2-methyl
104.071	$C_4H_6O_2$	0.7	dihydrofuran-3-one, 2-butenoic acid
106.087	$C_4H_8O_2$	0.1	2,5-Dimethyl-3-methylene p-dioxane
108.066	$C_3H_6O_3$	0.6	lactic acid, glyceraldehyde, dihydroxyacetone
114.056	$C_5H_4O_2$	1.3	Furfural, 2-pyran-2-one, 2H-Pyran-2-one, cyclopent-2-ene-1,4-dione
114.092	C_6H_8O	0.1	2,4-Dimethyl furan, 2-methylcyclopentenone
116.071	$C_5H_6O_2$	1.1	Furfuranol, 2-methyl-2(3H)-furanone, 5,6-dihydro-2H-pyran-2-one
118.087	$C_5H_8O_2$	0.3	2-propenoic acid, ethyl ester, 2-methylmethacrylate
120.066	$C_4H_6O_3$	0.9	3-Oxobutanoic acid, p-/o-dioxanone, 3-hydroxy-γ-butyrolactone
120.102	$C_5H_{10}O_2$	0.1	2H-pyran-4-ol, tetrahydro, tetrahydro-2-furanylmethanol
122.082	$C_4H_8O_3$	0.1	1,2-Ethanediol, monoacetate
128.071	$C_6H_6O_2$	0.4	1-(2-Furanyl)-ethanone, 5-methyl-2-furfural
130.087	$C_6H_8O_2$	1.0	Cyclotene
132.066	$C_5H_6O_3$	1.6	4-Hydroxy-5-methylfuran-3(2H)-one
132.102	$C_6H_{10}O_2$	0.4	2H-Pyran-2-methanol, 3,4-dihydro-, y-Caprolactone
134.082	$C_5H_8O_3$	0.8	Levulic acid
142.05	$C_6H_4O_3$	0.5	2,5-Furandicarbaldehyde, 2-oxo-2H-pyran-5-carbaldehyde
144.066	$C_6H_6O_3$	4.2	5-(Hydroxymethyl)-2-furfural, malton, levoglucosenone

Table 1 (continued).

m/z	Chemical Composition	Relative Abundance (% of Levoglucosan)	Tentative Chemical Compositions Assignments
144.102	$C_7H_{10}O_2$	0.2	2,6-Dimethyl-2,3-dihydro-4H-pyran-4-one
146.045	$C_5H_4O_4$	0.9	Aconic Acid, 5-hydroxy-2-furoic acid, rubiginol
146.082	$C_6H_8O_3$	0.4	a-Acetylbutyrolactone, 2,5-Bis(hydroxymethyl)furan
148.061	$C_5H_6O_4$	0.2	2-furancarboxylic acid, tetrahydro-5-oxo-
148.097	$C_6H_{10}O_3$	0.1	Pantoic lactone, 2,5-dimethoxy-2,5-dihydrofuran
150.077	$C_5H_8O_4$	0.8	2-deoxy-D-erythro-pentono-1,4-lactone
152.092	$C_5H_{10}O_4$	0.1	2-deoxy-α-D-erythro-pentopyranose
160.061	$C_6H_6O_4$	0.6	5-methoxy-2-furoic acid, 2-Furyl(hydroxy)acetic acid, kojic acid
160.097	$C_7H_{10}O_3$	0.1	2,2,6-Trimethyl-4H-1,3-dioxin-4-one
162.077	$C_6H_8O_4$	14.8	4H-Pyran-4-one, n,n'-dihydro-3,5-dihydroxy-n"-methyl, Anhydrolevoglucosan
164.092	$C_6H_{10}O_4$	0.6	D-Galactal
174.077	$C_7H_8O_4$	0.1	5-ethoxy-2-furoic acid, 5-(methoxymethyl)-2-furoic acid, 5-hydroxymethyl-2-furoic acid
176.092	$C_7H_{10}O_4$	0.4	Terebinic Acid, 2,2,5-trimethyl-1,3-dioxane-4,6-dione
178.072	$C_6H_8O_5$	1.6	Cortalcerone
180.086	$C_6H_{10}O_5$	100.0	Levoglucosan
182.103	$C_6H_{12}O_5$	3.1	2-Deoxy-d-Glucose
192.087	$C_7H_{10}O_5$	0.2	Shikimic Acid
208.082	$C_7H_{10}O_6$	0.5	3-Dehydroquinic acid
210.098	$C_7H_{12}O_6$	1.8	2,7-anhydro-β-D-altro-hept-2-ulopyranose, Quinic acid
212.113	$C_7H_{14}O_6$	0.1	Methyl α-D-glucopyranoside
222.097	$C_8H_{12}O_6$	2.4	5-(1,2-Dihydroxy-ethyl)-3,4-dimethoxy-5H-furan-2-one, gadusol
230.139	$C_{11}H_{16}O_4$	0.1	4-(Hydroxymethyl)-2-methyl-5-(2-methyl-2-propanyl)-3-furoic acid
240.108	C8H14O7	6.1	Unknown
324.129	C12H8O9	1.2	Unknown
342.14	$C_{12}H_{20}O_{10}$	9.2	Cellobiosan, β-D-Glucopyranose, 1,6-anhydro-4-O-β-D-galactopyranosy

The real-time monitoring capabilities of this technique also allow for monitoring metastable intermediates behavior over time. For example, the time profile of $[C_8H_{14}O_7+NH_4]^+$ (Figure 4c) has a peak maximum in-between the two curves of levoglucosan formation. We hypothesize that $C_8H_{14}O_7$ is a metastable intermediate that fragments down to levoglucosan ($C_6H_{10}O_5$) and acetic acid or hydroxyl-ethyl alcohol ($C_2H_2O_2$). The time profiles of *m/z* 222.097 and 324.129 also showed similar characteristics corresponding to metastable intermediates between cellobiosan and levoglucosan.

Time profiles for some furan-based compounds are shown in figure 4d and 4e. Furan compounds starts to appear approximately 1.0 second into pyrolysis while levoglucosan starts to appear 0.3 s after that. The appearance of furans before levoglucosan suggests that the formation of HMF and levoglucosan are effectively competing with each other, and HMF formation may be kinetically favored and levoglucosan may be thermodynamically favoured. This directly contradicts a previously suggested hypothesis that levoglucosan is an intermediate for furan formation[24] but supports Dauenhauer *et al* who suggests the two pathways are in competition.[8]

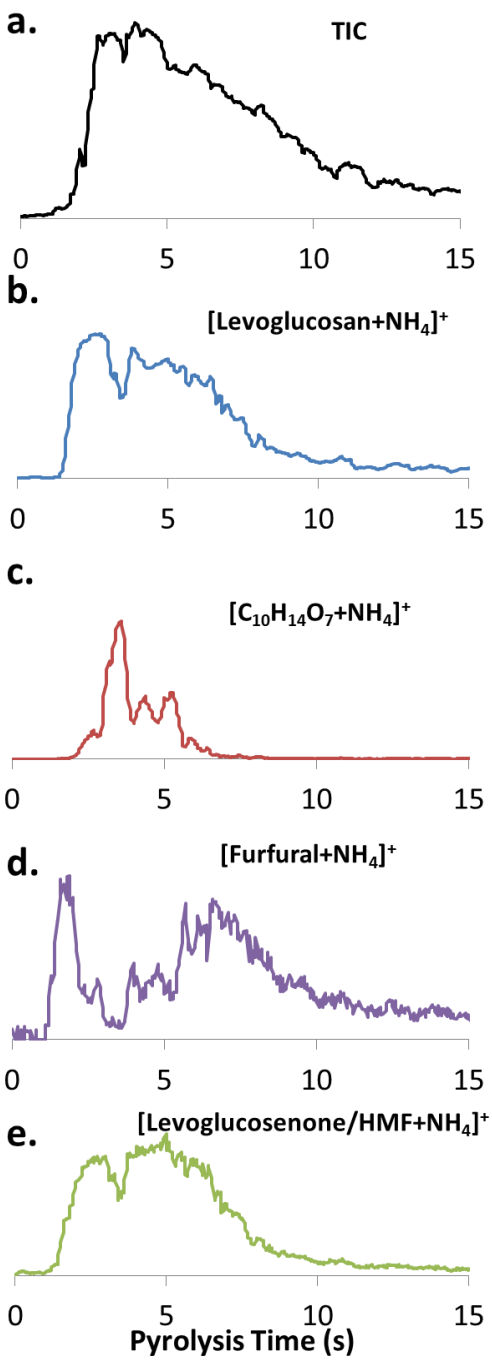


Figure 4. The total ion chromatogram (a) and extracted ion chromatograms for various pyrolyzates (b-d) obtain in the same pyrolysis run.

Powder vs. Thin-Film Cellulose Pyrolysis

Unfortunately, our data of temporal resolution in monitoring pyrolysis products cannot be directly translated into kinetic time profiles due to broadening of the temporal data. The broadening is especially problematic if the pyrolysis is conduction-limited. According to Dauenhauer and co-workers, in order to obtain isothermal kinetics the t cellulose thickness needs to be less than 10 μm at 500°C.[8] They demonstrated that thin-film (<10 μm) cellulose pyrolysis has different product yields compared to powder pyrolysis; this included a decrease in levoglucosan yield and increase in small oxygenates with thin-film cellulose pyrolysis. However, this study was performed by micropyrolysis followed by GC-MS analysis and cannot monitor the pyrolyzates in real-time or study unstable pyrolyzate intermediates that do not survive. According to SEM measurements (data not shown), the cellulose used in this study has a wide size distribution. In addition, about ten cellulose particles were stuck together when we loaded them into the pyrolysis that could not be separated out. Thus, our powder experiment is expected to not be isothermal and conduction or other effects are expected to affect the apparent chemical kinetics.

To further explore our capability of real-time monitoring at isothermal conditions thin-film cellulose pyrolysis was investigated. The time profile of levoglucosan from thin-film pyrolysis (figure 5b) shows two distinct profiles similar to powder pyrolysis (figure 5a) that is separated by a dip at around 3 second. However, the profile widths of the two curves are much narrower in thin-film cellulose pyrolysis (around 1 and 2 seconds, respectively) than in powder cellulose pyrolysis (around 2 and 4 seconds). We

believe these differences in pyrolysis width time are due to conduction-limitations during pyrolysis with a thicker biomass particle.

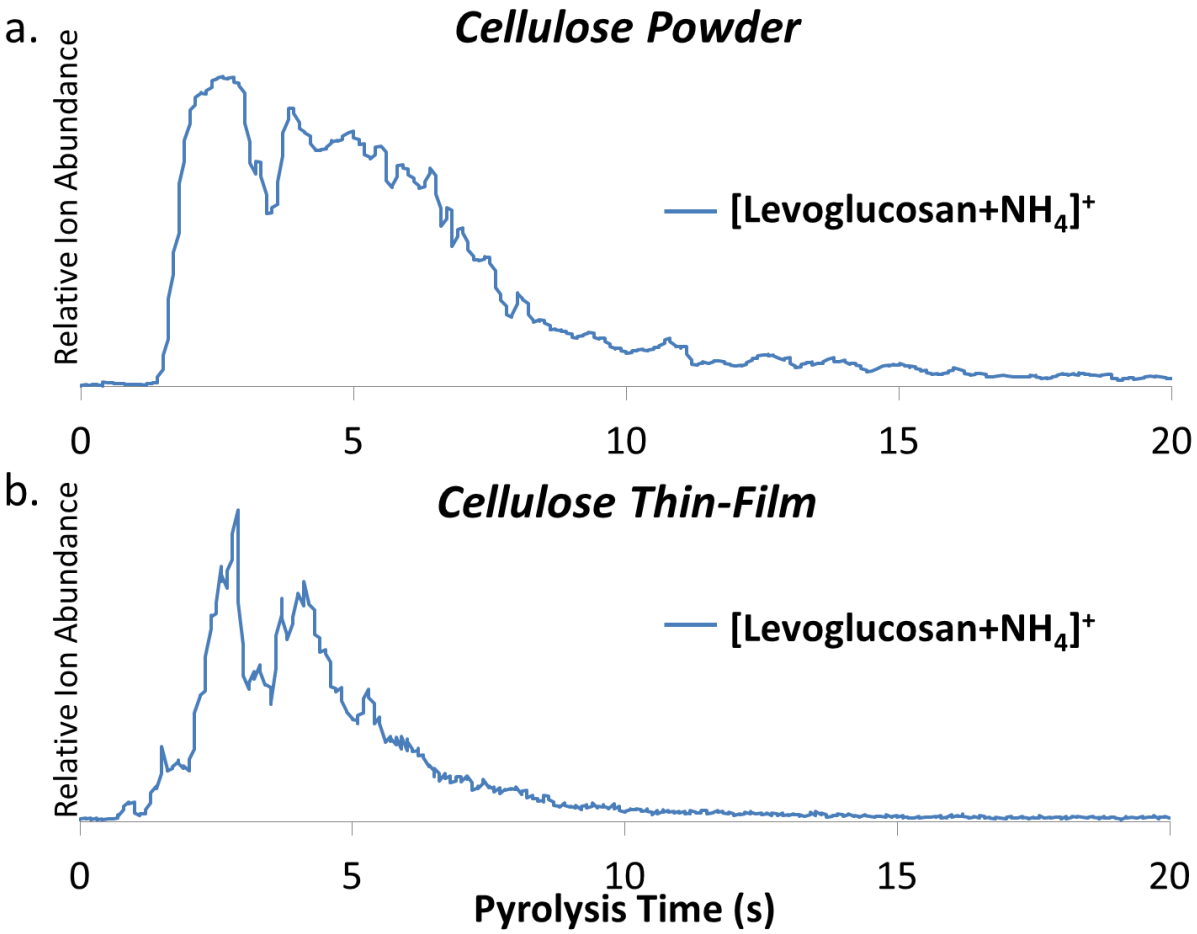


Figure 5. A comparison of levoglucosan's EIC between cellulose powder and cellulose thin film pyrolysis. Both the cellulose powder and thin-film pyrolysis show a distinct two-peak profile but the widths of these distributions differ. The increased width in the cellulose powder pyrolysis compared to the thin-film pyrolysis is attributed to a broadening effect of conduction-limited pyrolysis.

Pyrolysis of Single Red-Oak Particle

Pyrolysis of red oak biomass (composed of cellulose, hemicellulose, and lignin) is also studied using our µPy-APCI-TOF system. The main goal was to achieve a time-

resolved understanding of biomass pyrolysis that could be related to the pyrolysis in a pilot-scale or commercial-scale reactors. Therefore, the red oak used in this study was processed with the same procedure for biomass used in a pilot-scale reactor.[25] Because of the high sensitivity of the APCI-TOF, we could only pyrolyze a single red oak particle (filtered through a sieve size of ~ 200um, ~50µg). A single particle of red oak was pyrolyzed and he time profile of levoglucosan and a few furans (furfural, hydroxymethyl furfural and/or levoglucosenone) are shown in Figure 6a and 6b, respectively. The time profiles of red oak (Figure 6) show slightly different patterns from those of cellulose pyrolysis (Figure 4). For example, the time profile of levoglucosan in red oak has an initial spike in signal that appears within 100 ms of pyrolysis. The intense spike, or peak, is followed by two broader peaks that are similar to those in cellulose pyrolysis. The first peak's maximum intensity is approximately 1.5x that of the broader peaks, but the area is only 15% of the two broader peaks.

We hypothesize that the initial peak arises from the pyrolysis of hemicellulose; whereas, the later peaks are from cellulose (figure 6c). Hemicellulose is a branched biopolymer made of both pentose and hexose and is more loosely bound in the biomass structure with a much lower degree of polymerization. Therefore, we expect it can be more readily pyrolyzed. By comparison, cellulose is a strongly bound linear biopolymer that has a fibril structure created by hydrogen bonding between multiple cellulose monomer units.[26] The fibril structure is expected to require more time to heat up and pyrolyze compared to hemicellulose. In addition, cellulose is expected to take even more time to pyrolyze if it undergoes a molten phase transition before pyrolysis occurs

(~100 μs time scale at 700°C).[14]

Furfural and hydroxymethyl furfural (or levoglucosenone) show characteristics similar to the levoglucosan red oak time profile with an initial peak followed by two broad peaks. The two later peak profiles show similar behavior as those of cellulose pyrolysis: 1) they start to appear around one second after pyrolysis; 2) the two profiles are separated by a sharp dip at around 2 or 3 seconds; and 3) the ion abundances furans are about ten times less than that of levoglucosan. There exist minor differences but these are understandable considering the other factors such as the effect of alkaline metal in red oak and the differences in cellulose crystallinity and degrees of polymerization. The initial spike of furan signal is earlier in red oak than the initial peak of levoglucosan and has a much higher abundance than the late eluting, broader peak. The area of the initial peak for furfural and hydroxymethyl furfural (or levoglucosenone) is 50% and 75% of the area of the later peaks, respectively; this is a vast contrast to levoglucosan of which the initial peak area is only 5% of the area of the later peak. The dramatic increase of the early furans' signal further supports the hypothesis that early pyrolyzates (< 1 sec) are from hemicellulose. The major monosaccharide of hemicellulose is xylose (5-carbon ring sugar), which is known to preferentially dehydrate to furan compounds.[27]

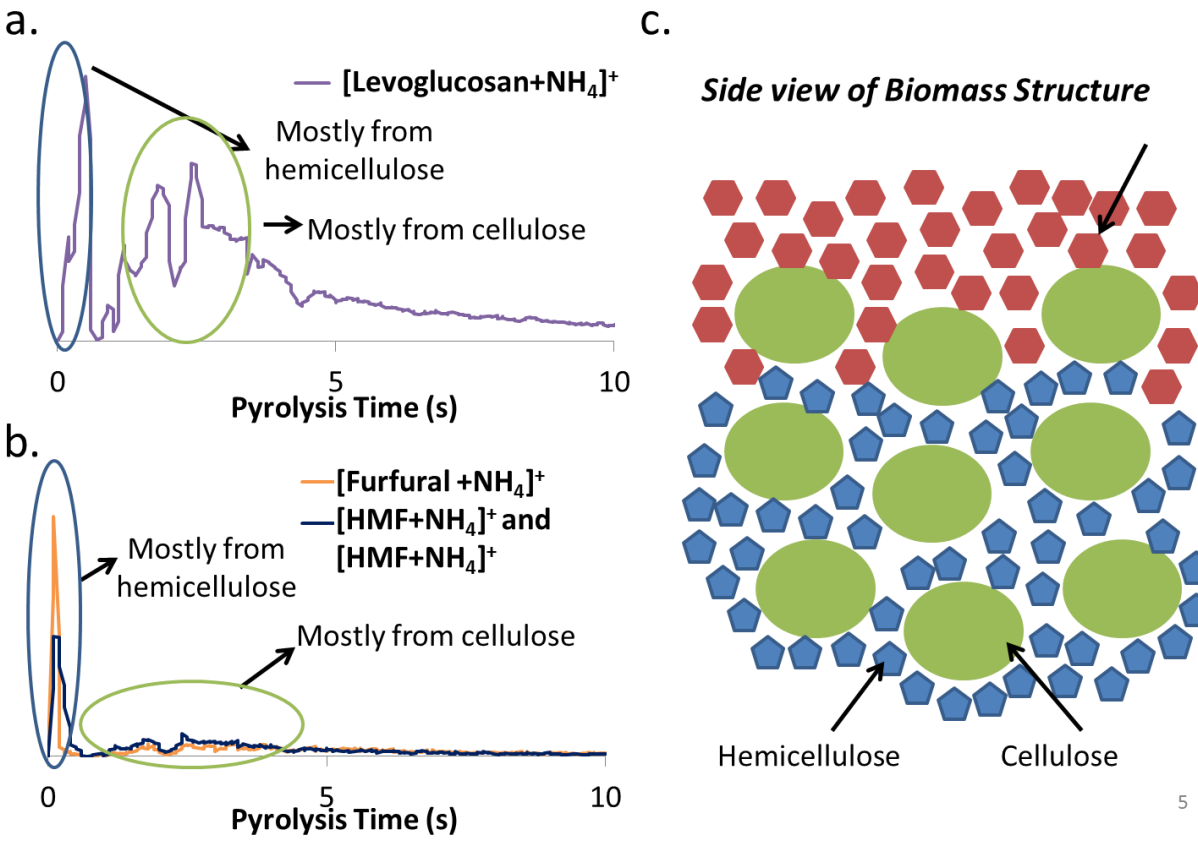


Figure 6. The pyrolysis of a single red oak particle revealed an additional peak of levoglucosan (a) and furan (b) signal that appears very rapidly and is no present in cellulose pyrolysis. This initial peak arises from the hemicellulose component of biomass that is less impeded in the biomass structure (c); and therefore, more readily starts pyrolyzing before cellulose.

The capability of the APCI-TOF to analyze single particle pyrolysis allows for the effect of particle heterogeneity on pyrolysis to be studied. For example, a dramatic difference was observed in the time profiles between flake-shaped and rod-shaped particles of red oak. Time profiles of a few selected red oak pyrolyzates from rod- (a) and flake–shaped (b) particles are shown in Figure 7. One of the most significant differences between the two shapes is in the levoglucosan time profile. As shown previously, a rod-shaped particle of red oak produces a levoglucosan time profile with an

initial spike in signal within 100 ms of pyrolysis that is followed by two broader peaks. However, when a flake of red oak is pyrolyzed, the profile of levoglucosan is quite different. There is still an initial spike in levoglucosan signal within 100 ms of pyrolysis, but this is followed by a gradual tailing off of the levoglucosan signal without the two broader peaks observed in the pyrolysis of the rod-shaped particle.

We attribute the difference in time profiles between the two particle shapes to differences in chemical makeup. The flake-shaped particle is believed to have higher hemicellulose content, as evident by the strong signal of pyrolyzates eluting within 100 ms. This is consistent with our previous observation that the initial peak of pyrolyzates arises from hemicellulose. The initial peak in the flake-shaped particle shows a dramatic increase in furans, like furfural and HMF, which is consistent with an increase in hemicellulose content. Additionally, Baxter and co-workers observed differences between a flake-like sawdust particle and a rod-like sawdust particle; they found that a flake-like particle lost mass/devolatilized slightly faster than a rod-shaped sawdust particle of approximately the same size and volume. [28] We believe this effect can be explained in our data by the increased production of furans from hemicellulose, which readily volatilize. The boiling points of furfural and HMF are approximately 116°C and 100°C, respectively;[29, 30] in comparison, the estimated boiling point of levoglucosan is 350°C.[31]

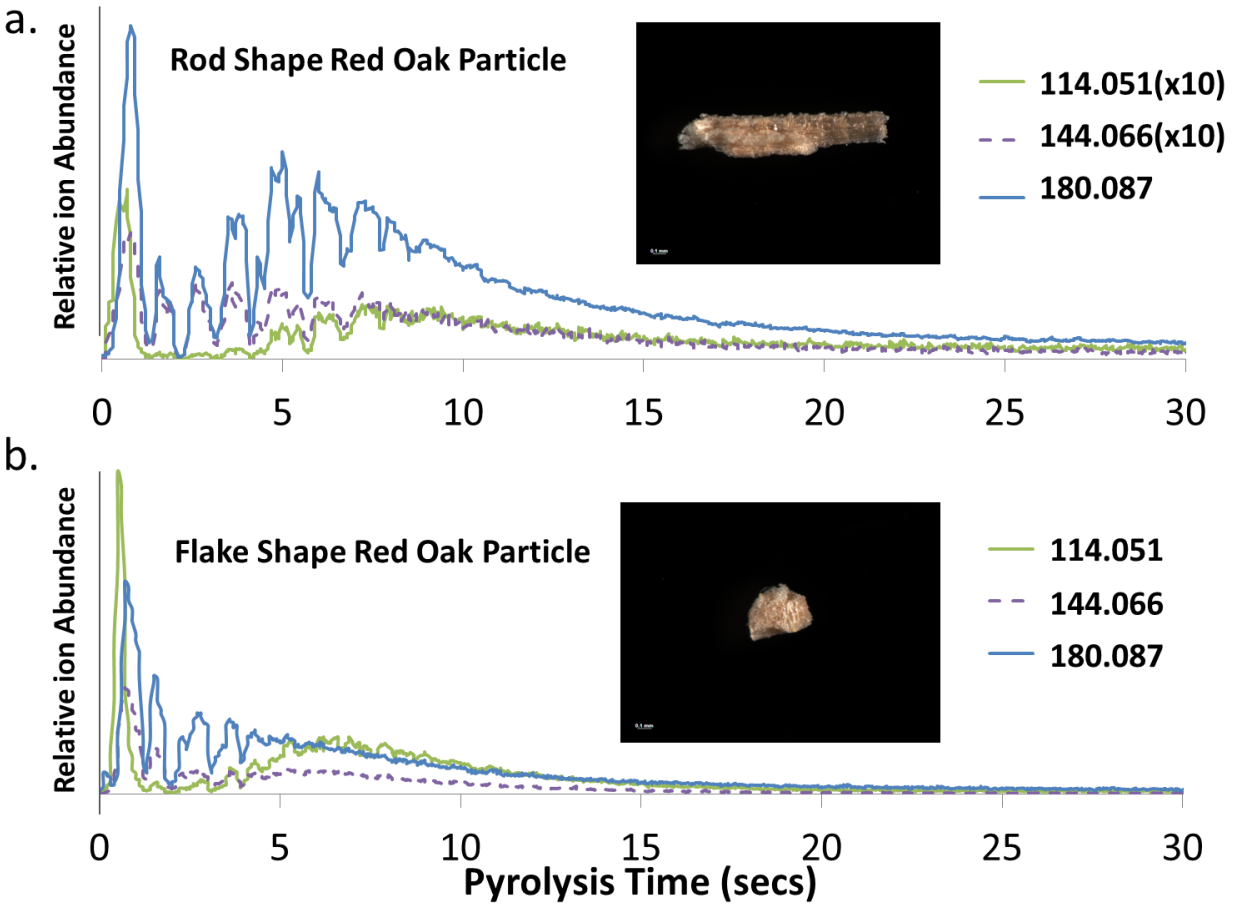


Figure 7. The sensitivity of the APCI-ToF allowed for single biomass particles to be studied; the extracted ion chronograms for a rod-shape red oak particle (a) is compared with a flake-shape red oak particle (b). The signals for furfural and HMF were multiplied by a factor of 10 in the rod-shape dataset (a).

Conclusion

In this work, the ability to monitor fast pyrolysis at the molecular level with μPy-APCI-TOF was explored. First, an equimolar mixture of model bio-oil compounds (like levoglucosan, ribose, and furans) showed that this technique is very sensitive and has a wide dynamic range. Also, the ability to observe only ammonium-adducted ions with minimal fragmentation allowed for the study of intact pyrolyzates rather than the

fragment ions that are traditionally observed in µPy-GC-MS studies. The high resolution of the TOF MS enabled the real-time analysis of pyrolyzates without GC separation. The ability to monitor pyrolyzates without a GC separation also allows for metastable intermediates of cellulose pyrolysis to be observed. These metastable intermediates are not seen with µPy-GC-APCI-ToF as they are thermally unstable and do not survive in the GC column. The temporal resolution curves (or EICs) also give evidence that these intermediates, like $C_8H_{14}O_7$, fragment down to levoglucosan. Furans start to appear shortly before levoglucosan, indicating that the formation of furans actually competes with the formation of levoglucosan.

The sensitivity of this analytical platform enables the study of single biomass particle pyrolysis. With this sensitivity, differences between shapes of red oak particles could be observed. For example, a flake-shape red oak particle produces furans more rapidly than a rod-shape particle; this is believed to be caused by differences in chemical make-up (differences in hemicellulose content). Also, fundamental differences on the effects of conduction and mass transport could be observed when studying a thick layer of cellulose powder (> 10µm) vs. a thin-film of cellulose (< 10µm). Unfortunately, the current state of these experiments does not allow for the time profiles of pyrolyzates to be directly converted into kinetic profiles. However, we feel with further study into the effects of conduction and mass transport on the EIC time profiles will allow for the correction of the observed broadening and will allow for kinetic information to be obtained.

Acknowledgments

This work was supported by Phillips 66 and Iowa State University. The authors thank Dr. Robert Brown, Patrick Johnston, and Dustin Dalluge at CSET from Iowa State University for providing the biomass as well as helpful discussions and useful advice during this project.

References

1. S. Czernik and A. Bridgwater, *Energy & Fuels*, 2004, **18**, 590-598.

2. S. Yaman, *Energy Conversion and Management*, 2004, **45**, 651-671.

3. D. Mohan, C. U. Pittman and P. H. Steele, *Energy & Fuels*, 2006, **20**, 848-889.

4. H. B. Mayes and L. J. Broadbelt, *The Journal of Physical Chemistry A*, 2012, **116**, 7098-7106.

5. A. Pollard, M. Rover and R. Brown, *Journal of Analytical and Applied Pyrolysis*, 2012, **93**, 129-138.

6. A. Broido and M. A. Nelson, *Combustion and Flame*, 1975, **24**, 263-268.

7. J. E. White, W. J. Catallo and B. L. Legendre, *Journal of Analytical and Applied Pyrolysis*, 2011, **91**, 1-33.

8. M. S. Mettler, S. H. Mushrif, A. D. Paulsen, A. D. Javadekar, D. G. Vlachos and P. J. Dauenhauer, *Energy & Environmental Science*, 2012, **5**, 5414-5424.

9. Q. Lu, X.-c. Yang, C.-q. Dong, Z.-f. Zhang, X.-m. Zhang and X.-f. Zhu, *Journal of Analytical and Applied Pyrolysis*, 2011, **92**, 430-438.

10. M.-K. Bahng, C. Mukarakate, D. J. Robichaud and M. R. Nimlos, *Analytica Chimica Acta*, 2009, **651**, 117-138.

11. H. Yang, R. Yan, H. Chen, D. H. Lee and C. Zheng, *Fuel*, 2007, **86**, 1781-1788.

12. O. Boutin, M. Ferrer and J. Lédé, *Journal of Analytical and Applied Pyrolysis*, 1998, **47**, 13-31.

13. P. J. Dauenhauer, J. L. Colby, C. M. Balonek, W. J. Suszynski and L. D. Schmidt, *Green Chemistry*, 2009, **11**, 1555-1561.

14. A. R. Teixeira, K. G. Mooney, J. S. Kruger, C. L. Williams, W. J. Suszynski, L. D. Schmidt, D. P. Schmidt and P. J. Dauenhauer, *Energy & Environmental Science*, 2011, **4**, 4306-4321.

15. J. Lédé, *Journal of Analytical and Applied Pyrolysis*, 2012, **94**, 17-32.

16. M. S. Mettler, D. G. Vlachos and P. J. Dauenhauer, *Energy & Environmental Science*, 2012, **5**, 7797-7809.

17. E. A. Smith and Y. J. Lee, *Energy & Fuels*, 2010, **24**, 5190-5198.

18. E. A. Smith, S. Park, A. T. Klein and Y. J. Lee, *Energy & Fuels*, 2012, **26**, 3796-3802.

19. D. Paul Cole, E. A. Smith, D. Dalluge, D. M. Wilson, E. A. Heaton, R. C. Brown and Y. Jin Lee, *Fuel*, 2013.

20. A. D. Paulsen, M. S. Mettler and P. J. Dauenhauer, *Energy & Fuels*, 2013, **27**, 2126-2134.

21. H. Leskinen, J. P. Suomela and H. Kallio, *Rapid communications in mass spectrometry*, 2007, **21**, 2361-2373.

22. C. J. Wachsmuth, M. F. Almstetter, M. C. Waldhier, M. A. Gruber, N. Nürnberger, P. J. Oefner and K. Dettmer, *Analytical Chemistry*, 2011, **83**, 7514-7522.

23. C. H. Stephens, P. M. Whitmore, H. R. Morris and M. E. Bier, *Biomacromolecules*, 2008, **9**, 1093-1099.

24. Y.-C. Lin, J. Cho, G. A. Tompsett, P. R. Westmoreland and G. W. Huber, *The Journal of Physical Chemistry C*, 2009, **113**, 20097-20107.

25. N. Kuzhiyil, D. Dalluge, X. Bai, K. H. Kim and R. C. Brown, *ChemSusChem*, 2012, **5**, 2228-2236.

26. R. C. Brown, *Biorenewable resources*, Iowa State Press, 2003.

27. J. N. Chheda, Y. Román-Leshkov and J. A. Dumesic, *Green Chemistry*, 2007, **9**, 342-350.

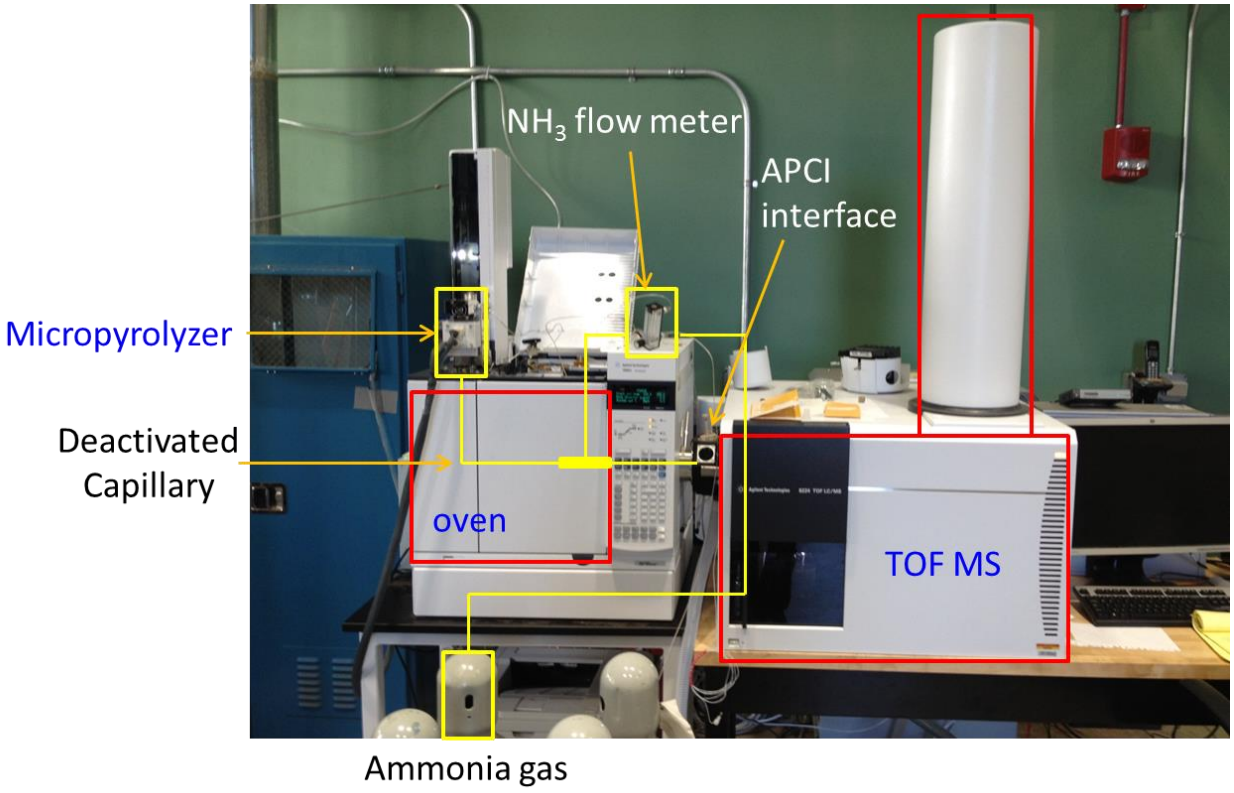
28. H. Lu, E. Ip, J. Scott, P. Foster, M. Vickers and L. L. Baxter, *Fuel*, 2010, **89**, 1156-1168.

29. J. Chickos and W. E. Acree Jr, *J. Phys. Chem. Ref. Data*, 2003, **32**, 519.

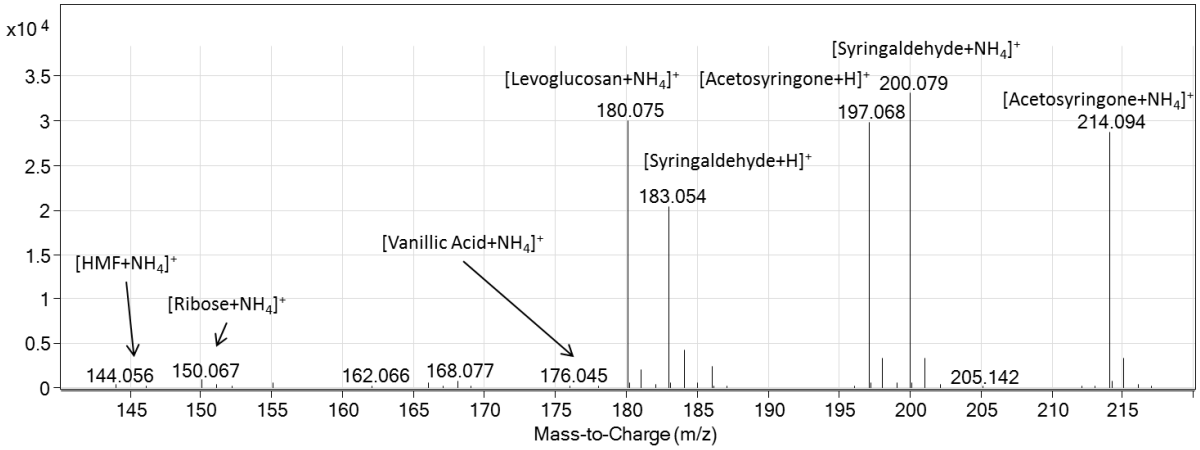
30. R.-J. van Putten, J. C. van der Waal, E. de Jong, C. B. Rasrendra, H. J. Heeres and J. G. de Vries, *Chemical reviews*, 2013, **113**, 1499-1597.

31. E. M. Suuberg and V. Oja, *Vapor pressures and heats of vaporization of primary coal tars*, Federal Energy Technology Center, Morgantown, WV (US); Federal Energy Technology Center, Pittsburgh, PA (US), 1997.

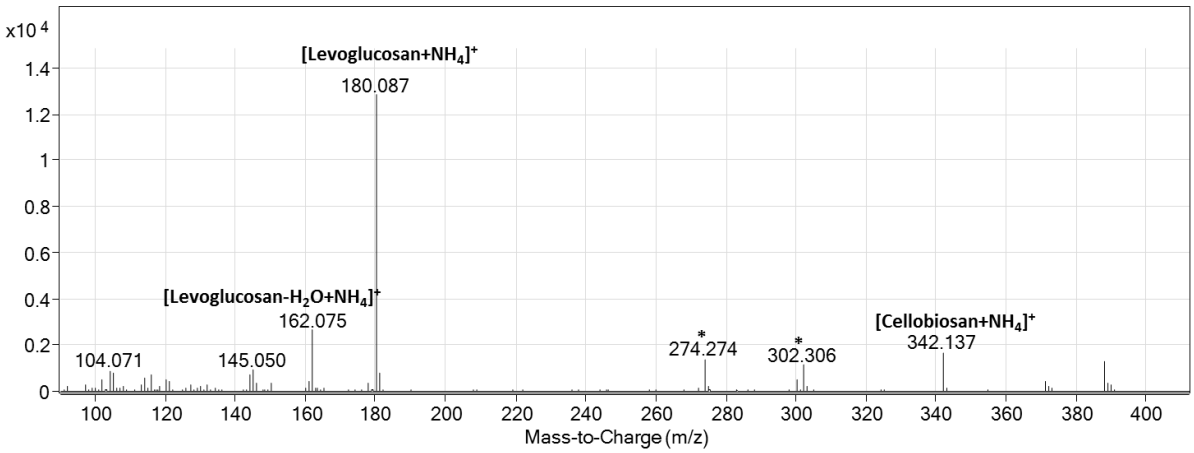
Supplementary Figure 1.



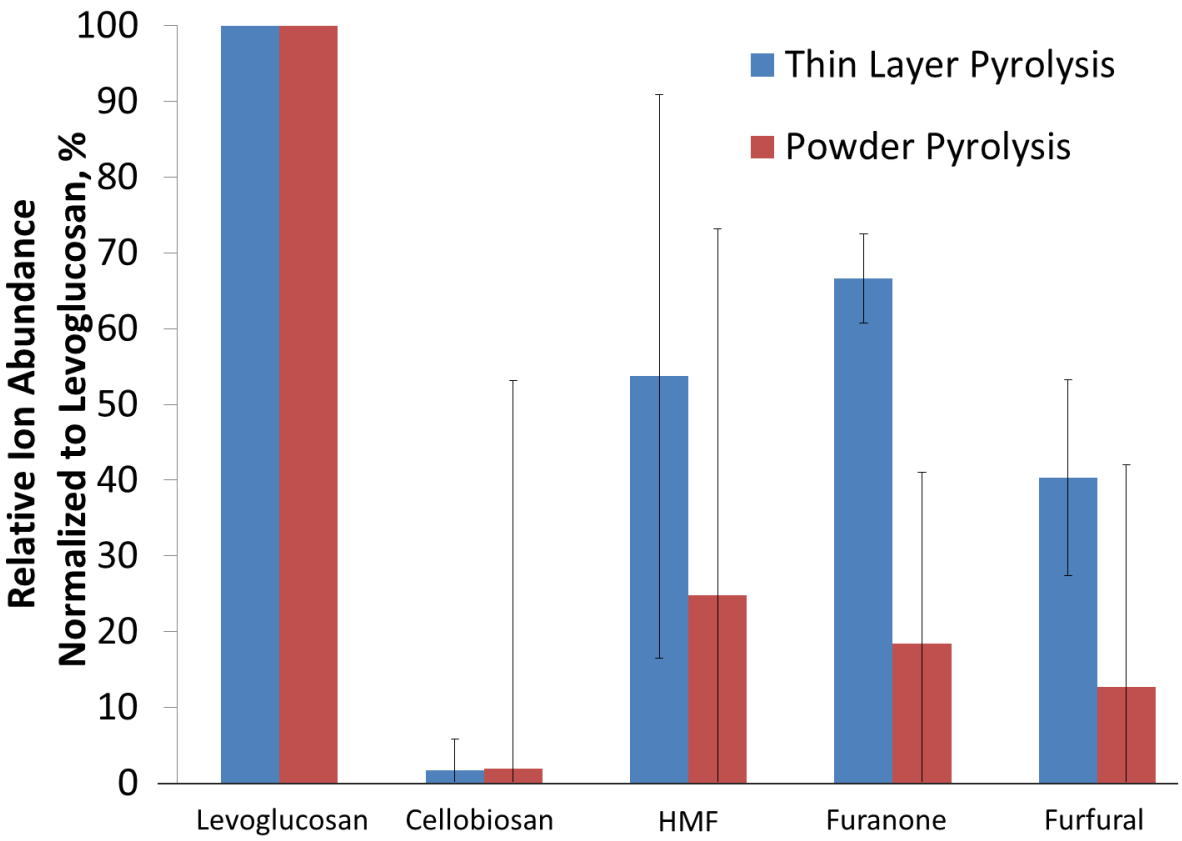
Supplementary Figure 2. Mass Spectrum of Standard Compounds



Supplementary Figure 3. Summed Mass Spectra for µPy-GC-APCI-TOF (J&W DB1701)



Supplementary Figure 4. Thin Layer Pyrolysis vs. Powder Cellulose Pyrolysis



CHAPTER VI

GENERAL CONCLUSIONS

Conclusions

The work presented in this dissertation has provided a solid foundation for using high resolution mass spectrometry (HRMS) to study biomass pyrolysis oils at the molecular-level. This work was successful in showing that HRMS, with its high resolving power and high mass accuracy, can identify hundreds of compounds in bio-oil. For example, the first-time attempt with laser desorption ionization (LDI)-HRMS identified over one-hundred non-volatile compounds in these oils. Further expanding to utilize a different ionization source, like negative mode-electrospray ionization (ESI), identified over eight-hundred volatile and non-volatile compounds in bio-oil. Most of these compounds are not readily characterized at the molecular-level with the traditional analytical techniques that are used to analyze bio-oil.

A 'petroleomic' adaptation to graphically represent the bio-oil has also been shown to be very useful. The heteroatom class distribution, double-bond equivalence (DBE) distribution, and carbon number vs. DBE allow for the construction of a comprehensive picture of the bio-oil. Also, the DBE distribution chart and DBE contour plot are able to provide great in-sight into whether the compounds are lignin-derived (DBE≥4) or holocellulose-derived (DBE<4) pyrolysis products.

Finally, a pioneering methodology that adapts the HRMS technique further to study fundamentals of biomass pyrolysis has been developed. The methodology is the

first of its kind and enables the real-time monitoring of biomass fast pyrolysis at the molecular-level. It was shown that the capabilities of this analytical platform can provide insights into biomass pyrolysis that have not been possible with traditional techniques. For example, metastable intermediates of cellulose pyrolysis could be identified and monitored with this novel approach. Also, fundamental pyrolysis studies, such as the effect of biomass shape and thickness, are possible with this technique due to the high sensitivity and time resolution of the time-of-flight mass spectrometer used.

Future Directions

As this work was the first in utilizing high resolution mass spectrometry to analyze biomass pyrolysis oils, the future applications and possibilities are endless. One aspect of this research that will continue to be revolutionary to biomass pyrolysis is the HRMS real-time monitoring technique. With further improvements and understanding of the biomass pyrolysis process, this novel technique will enable more than just molecular monitoring of pyrolysis but also allow for things like chemical kinetics and mechanisms to be studied at the molecular-level. And with these capabilities, the opportunities to study fundamental pyrolysis chemistry beyond bio-oil characterization could be pivotal in helping biomass pyrolysis oils to mature into a viable and exciting form of renewable energy for the future.

APPENDIX

PETROLEOMIC CHARACTERIZATION OF BIO-OIL AGING USING FOURIER-TRANSFORM ION CYCLOTRON RESONANCE MASS SPECTROMETRY

A paper published in *Bulletin of the Korean Chemical Society*

Erica A. Smith[1], Christopher Thompson[2], and Young-Jin Lee[1]*
[1]Department of Chemistry, Iowa State University, [2]Bruker Daltonics, Inc.

Abstract

Bio-oil instability, or aging, is a significant problem for the long-term storage of fast pyrolysis oils. We investigated bio-oil aging at the molecular level using Fourier-transform ion cyclotron resonance mass spectrometry. Petroleomic analysis suggests that bio-oil aging is resulted from the oligomerization of phenolic lignin products whereas 'sugaric' cellulose/hemicellulose products have negligible effect.

Introduction

An important advantage of fast pyrolysis oils or bio-oils, compared to other renewable energies such as solar or wind energy, is their storability as a liquid fuel; however, storage procedures currently in place for petroleum crude oils cannot be directly applied to bio-oils due to the significant difference in their chemical properties. These differences in chemical properties arise mostly from the high oxygen content of biomass pyrolysis oils.[1] The high oxygen content is also known to make bio-oil very reactive. During prolonged storage, the reactivity causes the bio-oil to form higher molecular-weight compounds, which increases the overall viscosity of the bio-oil.[2] The increase in viscosity and the corresponding decrease in volatility make the use of bio-oils in fuel

applications very problematic since these are undesirable characteristics in transportation fuels. [3]

Several approaches for stabilizing bio-oil have been proposed that focus on decreasing the viscosity of the aged bio-oil. One common practice is to preheat the oil before combustion to lower the viscosity.[4] Preheating bio-oil has been shown to accelerate polymerization reactions and to cause phase separation of the bio-oil. As a result, preheating the oil may lead to particulates clogging the fuel lines.[2] Another approach to dealing with the increased viscosity is to dilute the bio-oil with an alcohol. The bio-oil diluted in alcohol is known to slow down the polymerization reactions.[5] Unfortunately, the addition of alcohol to stabilize bio-oil is not economically feasible.

Currently, the most common procedure for studying bio-oil stability was developed by Diebold and Czernik.[5] This procedure involves 'rapidly aging bio-oil' by heating the bio-oil at a mild temperature (usually 90 °C) for a certain amount of time (usually 24 hours). Physical measurements, like viscosity and water content, are made before and after the rapid aging procedure. A round robin study of this methodology showed huge variation in the test results.[6] Oasmaa and co-workers claim that the inconsistencies in the round-robin study was probably caused by a lack of experience in the laboratories.[2]

Brown and co-workers have been coupling the traditional physical measurements with gel permeation chromatography (GPC) to further study bio-oil aging.[7] This has allowed for the increases in the molecular weight distribution observed by GPC to be correlated with an increase in viscosity that occurs as the bio-oil ages. Studies have also been

conducted that examine the bio-oil aging progression by dividing the 24 hour rapid aging test into 8 hour increments. They found that most of the aging occurs within the first 8 hour increment of the rapid aging test. This first 8 hour increment showed a sharp decrease in lower molecular weight compounds (<100 Da) and an increase in higher molecular weight compounds.[7]

Gel permeation chromatography provides the molecular weight distribution or degree of polymerization of bio-oil aging[7], but it cannot provide molecular information of the higher molecular weight compounds. A better understanding of the molecular constituents involved in bio-oil aging is expected to provide better insight into how to stabilize the bio-oil and/or slow down the aging process. Gas chromatography-mass spectrometry (GC-MS) has been widely utilized to study bio-oil,[8] but GC-MS is limited to analyzing only volatile and low molecular weight compounds of bio-oil. In addition, it cannot provide molecular information of the compounds not present in NIST EI-MS database. Hence, there is still a need for an analytical tool to examine bio-oil aging at molecular level, specifically for the higher molecular weight compounds. Marshall and coworkers have developed a high-resolution mass spectrometry (HRMS) approach to directly analyze chemical compositions of thousands of compounds in petroleum oils and understand their molecular characteristics, named as petroleomics[9]. Previously, we have reported the successful use of a petroleomics approach to characterize bio-oils at the molecular-level.[10-12] Here we apply this approach to bio-oil aging to understand the associated molecular changes. Electrospray ionization (ESI) in negative ion mode and atmospheric pressure photoionization (APPI) in positive ion mode were used for the

current study, which are widely used for the petroleum crude oil analysis.[13,14]

Experimental

Fast Pyrolysis and Rapid Aging

The bio-oil samples were provided by the Brown Group at Iowa State University. In short, the bio-oil was produced by fast pyrolysis of red oak with a pilot-scale fluidized bed reactor located at the Biocentury Research Farm at Iowa State University.[8] The samples were then subjected to an accelerated aging procedure by heating at 90 °C for 0, 8, 16, and 24 hours (T0, T1, T2, and T3). These increments have been shown to represent 0-12 months of naturally occurring aging of bio-oil at room temperature.[3] The rapid aging samples were diluted in methanol at a concentration of 1 mg mL^{-1} and stored at 4 °C until analysis. Nalgene bottles were used to store the bio-oils because of their chemical resistivity. The bio-oil samples were further diluted right before analysis to a concentration of 0.1 mg mL^{-1} in a solvent mixture that is appropriate for the specific ionization technique being used. The solvent mixture for (-) electrospray ionization (ESI) was 50% methanol in water and the solvent mixture for (+) atmospheric pressure photoionization (APPI) was 15% toluene in methanol.

Mass Spectrometry

The bio-oil samples were analyzed with two Fourier transform ion cyclotron resonance (FTICR) MS: 7T SolariX FTICR MS at Iowa State University for ESI experiments and 12T SolariX FTICR MS at Bruker facility in Billerica, MA, USA, for APPI experiments. ESI was run in negative-ion mode and APPI was operated in

positive-ion mode. Each instrument was carefully tuned for the mass range of interest and to minimize possible aggregation and/or fragmentation of the bio-oil compounds[11].

Data Analysis

The FTICR data was first calibrated with DataAnalysis software (Bruker) using known bio-oil peaks. The calibrated data was then imported to Composer (Sierra Analytics, Modesto, CA, USA) where it was further calibrated using a homologous series algorithm and assigned chemical compositions based on accurate mass. The mass accuracy of the assigned chemical compositions was limited to less than 3 ppm, and the relative abundance threshold for peaks being included in the analysis was 0.1%.

Results and Discussion

(+) Atmospheric Pressure Photoionization

The APPI-FTICR MS spectra were acquired in positive ion mode for three bio-oil aging samples (T1, T2, and T3) and a control (T0), and T0 and T3 spectra are compared in Figure 1. The *m/z* ranges observed in the two spectra are from *m/z* 150 to 900. Overall peak patterns are similar between the two with an average molecular mass slightly higher in T3. With a closer look, the high mass tail is much more prevalent in the T3 spectrum. For example, the peaks at *m/z* range of 500~580 in T3 has an equivalent relative abundance of that at *m/z* range of 430~500 in T0.

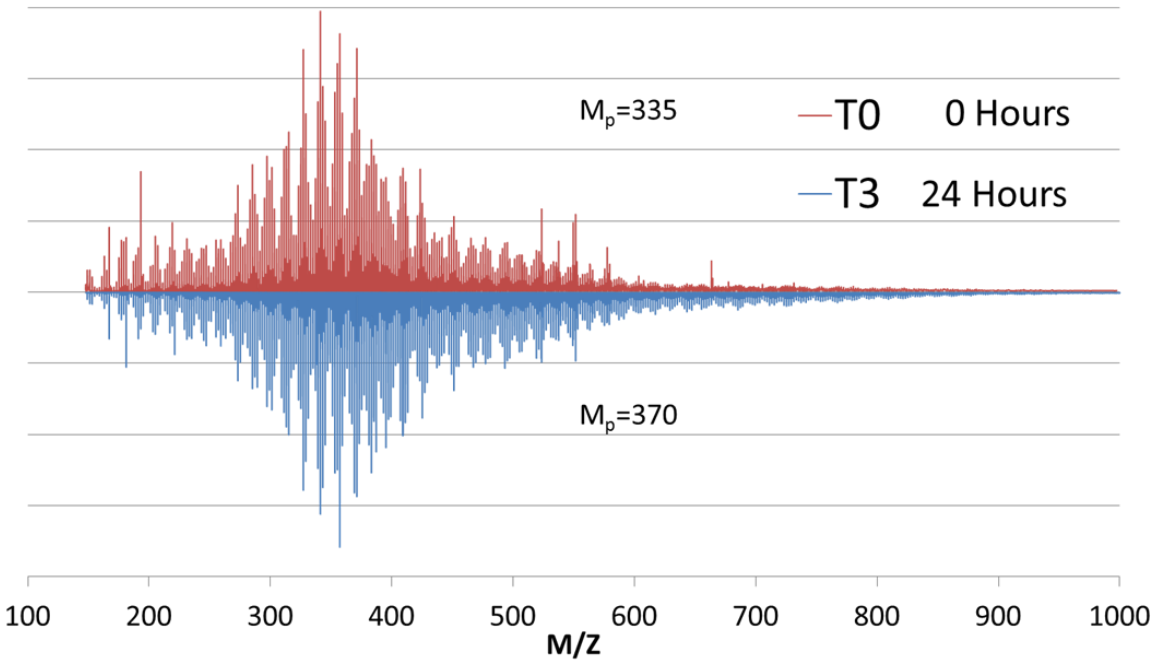


Figure 1. The (+) APPI mass spectra for un-aged bio-oil (T0, top red) and bio-oil that has been aged at 90 °C for 24 hours (T3, bottom blue).

Chemical composition analysis was performed for all four FTICR MS datasets. Both radical and protonated ions are produced by (+) APPI; however, there was no significant difference between radical and protonated ions in bio-oil samples in terms of overall heteroatom class distributions. Hence, the rest of the analysis was focused on protonated species because of their higher relative abundance in the presence of toluene dopant. The heteroatom class distributions for the four bio-oil aging samples are shown in Figure 2. The (+) APPI FTICR data of red oak bio-oil show only oxygen containing compounds, with the number of oxygen ranging from 0 (HC; hydrogen and carbon only) to 15. Compared to the bio-oil sample that did not undergo aging (T0), the ion abundance of aged samples (T1-T3) decreases in low oxygen compounds (O2-06) and increases in high oxygen compounds (O8 and higher) as the aging progress. This is in

good correlation with Figure 1 in the fact that high mass compounds have higher oxygen content.

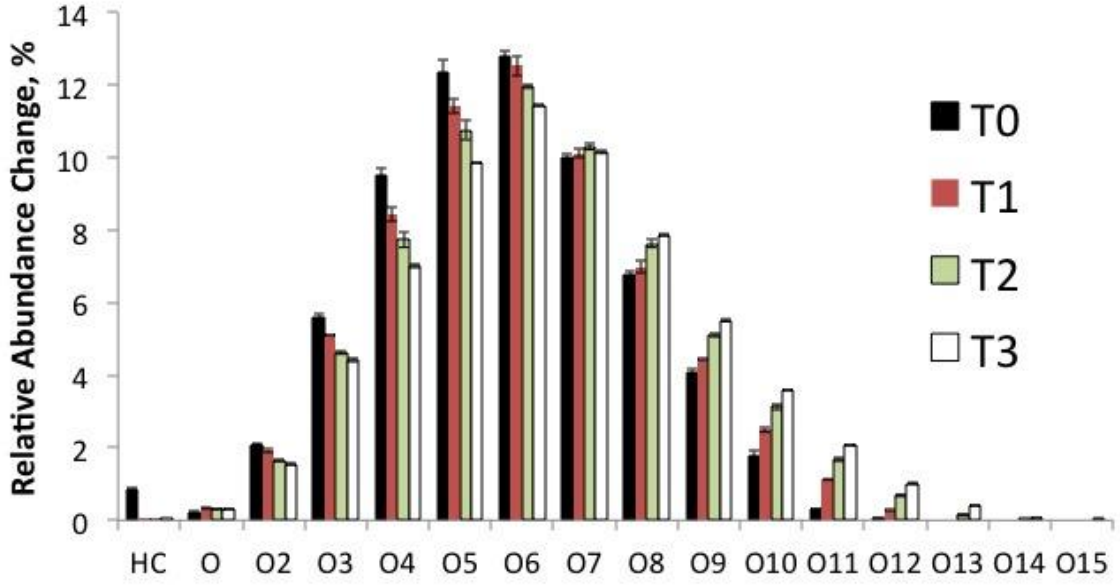


Figure 2. The protonated heteroatom class distribution of bio-oil aging samples obtained in (+) APPI-FTICR MS.

Double bond equivalence (DBE; the number of double bonds plus cyclic ring) is calculated from the following equation DBE = x - ½y + ½z + 1, with x, y and z being the number of carbon, hydrogen and nitrogen atoms, respectively. This is another useful tool in understanding complex bio-oil samples. APPI is well known to preferentially ionize aromatic compounds.[14] In case of bio-oils, phenolic compounds produced from lignin pyrolysis are the major compounds ionized by APPI. Each lignin monomeric unit has a minimum DBE value of four (i.e. benzene ring) and average DBE of about five (including one carbonyl or vinyl side chain).[10] Hence, we can estimate the degree of oligomerization in lignin pyrolysis products by dividing their DBE values by 5.

The DBE distribution of a few major heteroatom classes is shown in Figure 3 for T0 and T3 bio-oils. The overall distribution is similar for O4 class between T0 and T3 samples, specifically most abundant at DBE of 12 (dimer or trimer) and widely distributed over the range of 5-20 (monomer to tetramer); however, the relative abundance of the O4 class is decreased by 20~25% for T3 compared to T0. The O12 class is present in almost negligible amount for T0 but significant in T3 and most abundant at DBE of 20-23 (tetramer or pentamer). The O8 class compound is quite interesting for having two distinct distributions, one peak at DBE of ~12 and the other at ~16. The O8 relative abundance is increased by two-fold with T3 compared to T0. This data clearly suggests that the decrease of smaller oligomers (dimer or trimer) and the increase of bigger oligomers (tetramer or higher) as bio-oil aging.

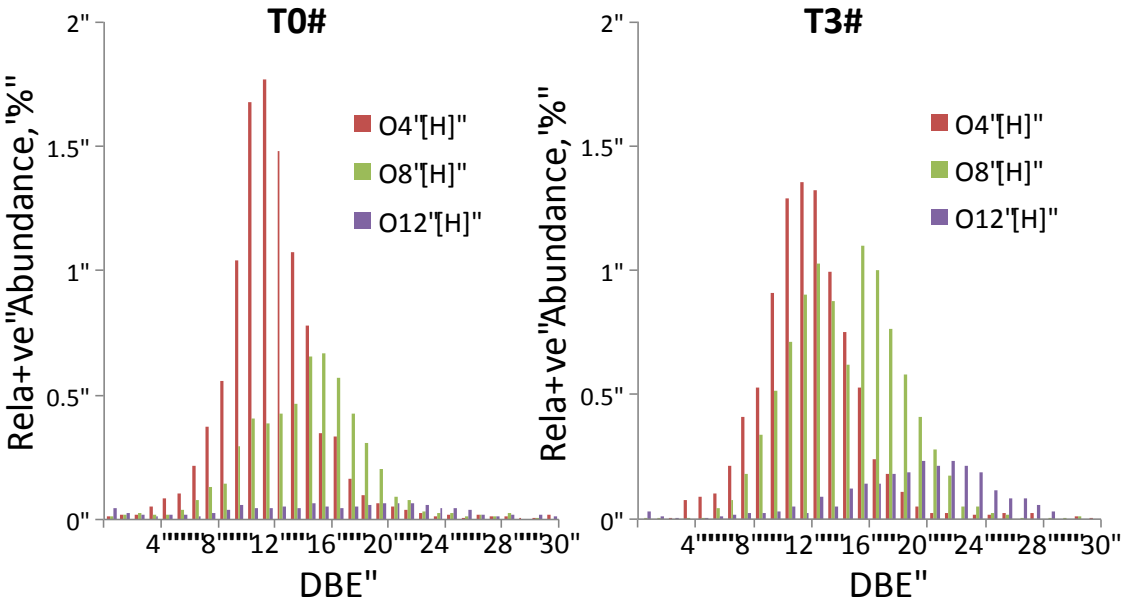


Figure 3. A DBE distribution comparison of T0 and T3 for a few major heteroatom classes observed in (+) APPI.

(-) Electrospray Ionization

To examine the effects of polar compounds on bio-oil aging, the bio-oil aging samples were analyzed using negative-mode electrospray ionization. (-) ESI has been previously shown to readily ionize both aliphatic (or sugar) and aromatic (or phenolic) compounds of bio-oil as deprotonated ions, $[M-H]^-$.[11] The spectra obtained for T0 (blue, top) and T3 (red, bottom) samples, in Figure 4, show only a slight increase of high mass compounds.

As we have previously reported [11], we can observe both cellulose/hemicellulose pyrolysis products and lignin pyrolysis products in (-) ESI. They can be easily distinguished from the difference in their DBE values. We separated the two heteroatom class distributions as shown in Figure 5. Here we define "sugaric" compounds as those with DBEs of three or less and "phenolic" compounds as those with DBE of four or higher. "Sugaric" compounds are cellulose and hemicellulose derived pyrolysis products like levoglucosan and "phenolic" compounds are from lignin pyrolysis[11]. Furans like hydroxymethyl furfural are five-membered aromatic compounds with a DBE of 3 and indistinguishable from some compounds like levoglucosenone (DBE of 3). In any case, they could still be counted as "sugaric" compounds because they are pyrolysis products of hemicellulose and cellulose. It should be noted this classification has some limitations as some of them might be overlapping. For example, furan or levoglucosenone with an additional double bond (from carbonyl or alkenyl side chain) will have a DBE value of 4 and counted as "phenolic". However, such contribution is expected to be minimal.

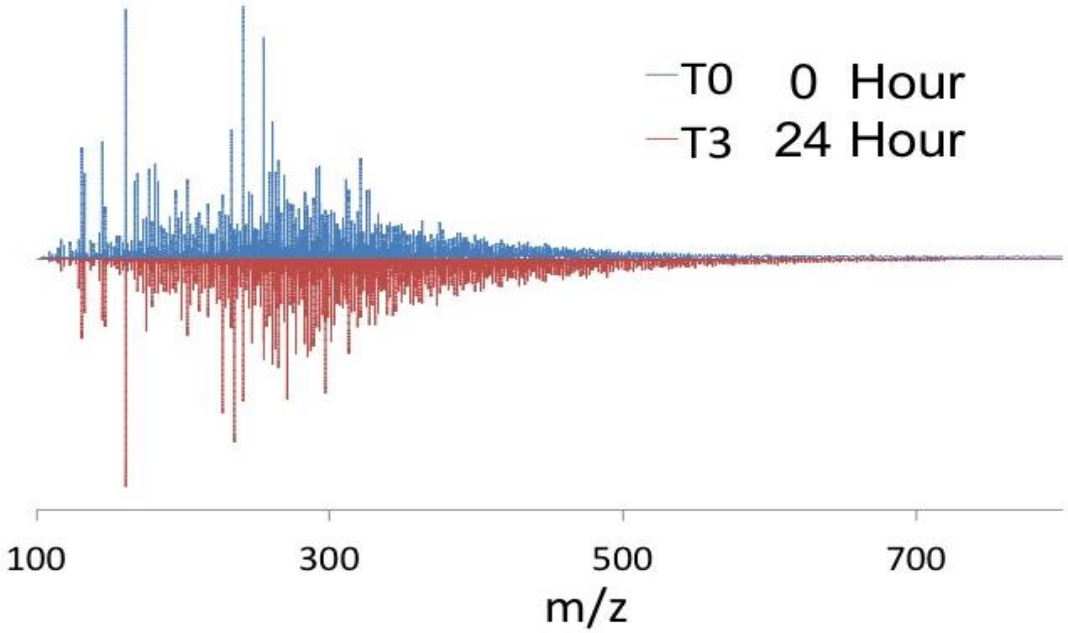


Figure 4. (-) ESI mass spectra for un-aged bio-oil (T0, top red) and after aging at 90 °C for 24 hr (T3, bottom blue).

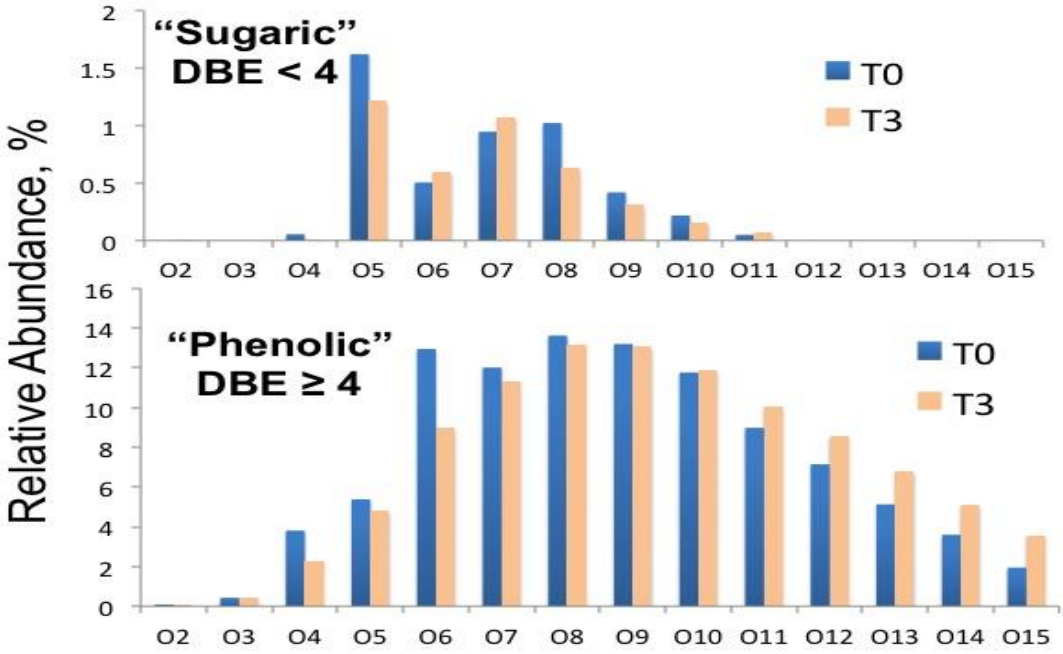


Figure 5. Heteroatom class distributions for "sugaric" (DBE<4) and "phenolic" (DBE ≥ 4) compounds for T0 and T3 samples.

The heteroatom class distribution for "sugaric" compounds (Figure 5, top) is similar between the two bio-oil aged samples. The slight difference in relative abundance for O5 and O8 compounds is attributed to pH matrix effects. We have previously reported that the relative abundance of levoglucosan (anhydrous glucose, $C_6H_{10}O_5$, O5 with DBE of 2; *m/z* 161, the highest abundance peak in Figure 4) is subject to pH matrix effects and shows the decrease of ion abundance at low pH.[11] It should be noted there is no apparent difference for the *m/z* 161 peak in Figure 4 because the spectra are normalized to the highest abundance peak (*m/z* 161), but the relative ion abundance of *m/z* 161 is ~25% less for T3 when the spectra are normalized to the total ion count.

The aging effect is more apparent in the heteroatom class distribution of aromatic compounds (Figure 5, bottom) and shows a similar trend with (+) APPI data: a decrease of lower oxygen compounds and increase of higher oxygen compounds. This suggests that bio-oil aging mostly arises from the oligomerization of lignin pyrolysis products. Several possible aging mechanisms have been proposed including oxidation-induced reactions and esterifications.[4] Among those, acid catalyzed reactions of phenolic compounds is a possible mechanism for the oligomerization of lignin products.

Phenolic compounds observed in (-) ESI and (+) APPI have slightly different chemical functionalities. Phenolic compounds in (-) ESI have on average two or three more oxygens than in (+) APPI with the most abundant heteroatom class of O8/O9. This distribution extends beyond O15 (O16 and higher heteroatom classes are not shown) whereas in (+) APPI most abundant heteroatom class is O6 and very minimal for heteroatom classes above O14. This difference arises from their difference in ionization

efficiencies; (-) ESI preferentially ionizes compounds that can be readily deprotonated whereas (+) APPI preferentially ionizes compounds that can be readily protonated. While most phenolic compounds can be ionized by both modes, some compounds show a much higher abundance in one mode versus another. For example, phenolic compounds with a carboxylic group, e.g. vanillic acid, have a much higher abundance in (-) ESI. The fact that both (+) APPI and (-) ESI support the oligomerization of phenolic compounds in spite of the difference in ionization preference indicates that bio-oil aging occurs in a wide class of lignin pyrolysis products.

Conclusion

We have performed a petroleomic analysis of bio-oil aging utilizing high-resolution FTICR. We conclude that the problems with bio-oil stability arise mostly from the oligomerization of lignin derived pyrolysis products. This is supported by both (+) APPI and (-) ESI data for a wide class of lignin pyrolysis products. In contrast, "sugaric" compounds in (-) ESI (DBE < 4), specifically cellulose and hemicellulose pyrolysis products, show almost a negligible effect by aging.

It should be noted there are several limitations in the current analysis. Petroleomic analysis are not amenable to characterize low mass compounds (i.e. MW <100) whereas the change in pH should mostly arise from low oxygenates such as formic and acetic acids. We have observed dramatic change of a few ion peaks in (-) ESI with aging (Figure 4); however, we did not pursue detailed analysis of their origin because it is partially due to the pH matrix effect and is difficult to calibrate for without

the exact knowledge of their identities. Nevertheless, our study demonstrated the molecular weight increase in bio-oil aging is most likely due to the oligomerization of phenolic compounds from lignin pyrolysis.

Acknowledgments

This study is supported by a grant from Phillips 66. The authors are grateful for the bio-oil samples and useful discussions provided by the Brown group, and for the Composer program kindly provided by David Stranz, Sierra Analytics. EAS acknowledges partial support from Graduate Assistance in Areas of National Need (GAANN) fellowship from U.S. Department of Education.

References

(1) Czernik, S.; Bridgwater, A. V. *Energy Fuels*, **2004**, *18*, 590.

(2) Oasmaa, A.; Korhonen, J.; Kuoppala, E. *Energy Fuels*, 2011, **25**, 3307-3313.

(3) Czernik, S., Storage of biomass pyrolysis oils, In *Proceed. Biomass Pyrolysis Oil Properties Combustion Meeting*, Nat. Ren. Energy Lab, Golden, CO, U.S.A. 1994; p. 67.

(4) Diebold, J. P. In *Fast Pyrolysis of Biomass: A Handbook* Bridgwater, A. V., Ed.; CPL Press, Liberty House, U. K., 2002; p. 243.

(5) Diebold J. P.; Czernik, S. *Energy Fuels*, **1997**, *11*, 1081.

(6) Oasmaa A.; Meier, D. *J. Anal. Appl. Pyrol.*, **2005**, *73*, 323.

(7) Brown, R. C., A Systems Approach to Bio-oil Stabilization, 2011 OBP Program Review Portal, US-DOE, 17 Feb 2011, http://obpreview2011.govtools.us/.

(8) Pollard, A.; Rover, M.; Brown, R. C. *J. Anal. Appl. Pyrol.*, **2012**, *93*, 129.

(9) Marshall, A.; Rodgers, R. P. *Acc. Chem. Res.* **2004**, *37*, 53.

(10) Smith E. A.; Lee, Y. J. *Energy Fuels*, **2010**, *24*, 5190.

(11) Smith, E. A.; Park, S.; Klein A. T.; Lee, Y. J. *Energy Fuels*, **2012**, *26*, 3796.

(12) Cole, D. P.; Smith, E. A.; Dalluge, D.; Wilson, D. M.; Heaton, E. A.; Brown, R. C.; Lee, Y. J. *Fuel*, **2013**, *111*, 718.

(13) Yeo, I.; Lee, J. W.; Kim, S. Bull. Kor. Chem. Soc. **2010**, *31*, 3151-3155.

(14) Ahmed, A.; Ghosh, M. K.; Choi, M. C.; Choi, C. H.; Kim, S. J. Am. Soc. Mass Spectrom. **2013**, *24*, 316.