ABSTRACT

Advancing Ion Mobility – Mass Spectrometry Methods for the Analysis of Compositionally Complex Mixtures

Birendra Dhungana, Ph.D.

Mentor: C. Kevin Chambliss, Ph.D.

Enormous analytical challenges are involved in the analysis of compositionally complex samples such as biomass-derived products and crude oils. In this dissertation mass spectrometry (MS)-based novel strategies for the characterization and identification of chemical components in pyrolysis bio-oils are presented. Several techniques including, accurate mass analysis (*e.g.*, Kendrick mass defect analyses, chemical formula determinations, and van Krevelen plots), collision-induced dissociation (CID) and mass – mobility relationships are utilized to support molecular-level investigations.

Oily and aqueous fractions of bio-oils produced by slow pyrolysis of two feedstocks, pine shavings (PS) and corn stover (CS), were analyzed by negative-mode electrospray ionization-Orbitrap and ion mobility-time-of-flight (IM-TOF) MS. Oxygen-rich species with a high degree of unsaturation were observed, indicating that catalytic upgrading will likely be required if slow-pyrolysis bio-oils are to be utilized as fuel. Additionally, results from analyses of the PS oily fraction subjected to a sequential solvent fractionation demonstrated a partial separation and enrichment of compounds

according to oxygen classes, where oxygen classes generally trended with solvent polarities. Generally, higher oxygen classes were preferentially enriched in higher polarity solvents.

Mass – mobility correlations were investigated by IM-MS for various structurally-unique homologous series composed of commercially-available compounds. Structural variation involved the inclusion of different repeat units in oligomeric series and different terminal groups in CH_2-homologous series. Mass – mobility correlations were also investigated for select CH_2-homologous series identified in a bio-oil and compared with results observed for commercial series. A linear mass – mobility correlation ($R^2 \geq 0.996$) was established for all series except those in which a substantial change in the gas-phase conformation of ions was probable. Slopes observed for CH_2-homologous series with a single terminal group were significantly steeper than slopes observed for series containing two terminal groups. Additionally, a correlation between slope and double bond equivalents suggested that the CH_2-homologous series identified in the bio-oil were structurally similar to commercial series containing two terminal groups. Additionally, IM-MS and IM-MS/MS analysis of ions belonging to select CH_2-homologous series suggested that mass – mobility correlations and post-ion mobility CID mass spectra may be useful in defining structural relationships among members of a given Kendrick mass defect series.

Advancing Ion Mobility - Mass Spectrometry Methods
for the Analysis of Compositionally Complex Mixtures

by
Birendra Dhungana, B.Sc., M.Sc.

A Dissertation

Approved by the Department of Chemistry and Biochemistry

_____
Patrick J. Farmer, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy

Approved by the Dissertation Committee

_____
C. Kevin Chambliss, Ph.D., Chairperson

_____
Touradj Solouki, Ph.D.

_____
Elyssia Gallagher, Ph.D.

_____
Patrick J. Farmer, Ph.D.

_____
William C. Hockaday, Ph.D.

Accepted by the Graduate School
May 2016

_____
J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	*Description*
2D	two dimensional
APCI	atmospheric pressure chemical ionization
APPI	Atmospheric pressure photoionization
CEM	charge ejection model
CID	collision-induced dissociation
CRM	charge residue model
CS	corn stover
CSAF	corn stover aqueous fraction
CSOF	corn stover oily fraction
DBE	double bond equivalents
DC	direct current
DCM	dichloromethane
ESI	electrospray ionization
FT-ICR	Fourier transform ion cyclotron resonance
GC-MS	gas chromatography – mass spectrometry
GPC	gel permeation chromatography
HRMS	high resolution mass spectrometry
IEM	ion ejection model
IM	ion mobility
KMD	Kendrick mass defect
LDI	laser desorption/ionization
m/z	mass-to-charge ratio
MALDI	matrix-assisted laser desorption/ionization
MD^2	second-order mass analysis

Abbreviation	*Description*
MS	mass spectrometry
NOM	natural organic matter
PEG	polyethyl glycol
PPG	polypropyl glycol
PS	pine shavings
PSAF	pine shavings aqueous fraction
PSOF	pine shavings oily fraction
RF	radio frequency
SDS	sodium dodecyl sulfate
TG	thermogravemetry
TOF	time of flight
TWIMS	travelling wave ion mobility spectrometry
Ver.	version

ACKNOWLEDGMENTS

I am deepely appreciative of many individuals who have supported me during my graduate study. Without their time, support, encouragement, I would not have completed this dissertation.

I thank my advisor Dr. Kevin Chambliss for your priceless support, guidance, encouragement and patience over the last several years. I am equally grateful to Dr. Christopher Becker for your guidance, patience and thoutful feedback on my research work. I truly believe that I would not be where I am today if two wonderful people were not there to guiding me through. Thank you for believing in me more than myself.

I am grateful to other faculty members, particularly Dr. Touradj Solouki and Dr. William Hockaday for invaluable suggestions and guidance. I also thank Dr. Patrick Farmer, Dr. Sascha Usenko and Dr. Elyssia Gallagher for serving on my committee.
I am thankful to graduate program director Dr. Charles Garner, Department of Chemistry and Biochemistry and Baylor University.

I thank Dr. Dennis Rabbe for your help and support in teaching laboratory. I am also thankful to Dr. Alejandro Ramirez for your technical assistance.

I thank many individulas in Chambliss and Solouki research groups including, Chrish Rash, Dr. Richard Sevcik, Dr. Bowen Du, Dr. Daniel Olaitan, Dr. Behrooz Zekavat, Dr. Brett Harper and Matt Brantley for assising me with research/proof-reading dissertation. I particularly thank Dr. Sharon Munisamy for your help and support during

early years of graduate study. I greatly appreciate your thoughtful feedback in my research work and writing.

I thank all Nepalese, residing currently or in the past, in Waco for your support and providing home-like environment when I am thousands of miles away from my homeland. It would have been much difficult to accomplish this if you were not there. Particularly, I am deeply grateful to Dr. Basu Pandthi and Sharmila Gautam-Gyawali for your incommensurable help during my first year of my graduate study.

Last but not least, I thank my loving and caring family, my wife, Sudha, and two kids, Sabrina and Sabal. Thank you to my parents for your continuous encouragement and selfless support.

DEDICATION

To my late father Kedar Nath Dhungana

CHAPTER ONE

Introduction to Pyrolysis Bio-oil and Relevant Analytical Techniques

Pyrolysis of Biomass

Currently, the world depends on fossil fuels to supply nearly 80% of its energy needs.[1] The continuous depletion of resources, as well as the adverse impact of burning fossil fuels on the environment, have been topics of concern in recent decades, leading to growing interest in alternative and renewable resources of energy. The potential of biomass has long been considered as a sustainable energy source due to its large abundance and carbon neutrality (*i.e.*, producing and using biomass derived energy, specifically utilizing biomass from fast growing crops, adds no net carbon dioxide to the environment).[2–5] Biomasses are organic materials derived from living or recently living organisms such as wood, waste materials, *etc*. Billions of tons of plant biomass waste are generated every year around the world, and it is presumed that efficient utilization of biomass can substantially reduce the world's dependence on fossil fuel. Accordingly, research focused on developing suitable techniques for the optimal utilization of biomass has surged considerably in the past few decades.[2,6–9]

Pyrolysis is a thermal treatment of organic material at high temperatures (> 300 ^{o}C) in the absence of oxygen.[10] Pyrolysis of biomass produces numerous decomposition products of lignocellulosic macromolecules in the form of pyrolysis vapors, aerosols and solid residue.[10,11] The liquid obtained by the condensation of pyrolysis vapors and aerosols is called bio-oil, and it is an attractive feedstock for fuel and/or value-added

chemicals.[5,12,13] Nevertheless, it is important to note that bio-oil is a complex mixture consisting of varieties of oxygen-rich and highly polar compounds that are produced as a result of many simultaneous and sequential reactions.[14–16] Bio-oil requires suitable upgrading (*e.g., via* de-oxygenation) before it can be used efficiently as fuel.[7] Moreover, before such strategies can be optimized, the molecular-level understanding of bio-oil composition is essential. To this end, compositional complexity of pyrolysis bio-oil poses substantial analytical challenges.

Traditional Techniques of Bio-oil Analysis

Many of the early studies that characterized bio-oils have measured acidity, elemental composition, water content, solubility, viscosity, density, heating values, *etc*.[17,18] Results from these studies suggest that bio-oils are highly acidic (pH ~2.5), highly viscous (viscosity ~30-200 centipoise) and dense (specific gravity ~1.2) liquids with heating values about one half those typical of fossil fuels. These bulk properties are important assays of crude bio-oil quality but give little of the molecular information needed to determine the upgrade potential for fuel and/or chemical resources.

Results from studies involving Fourier transform infrared spectroscopy,[3,18–20] nuclear magnetic resonance,[17,18,21,22] and gas chromatography-mass spectrometry (GC-MS)[11,15,19,23–26] have provided additional details about chemical structure. The two former techniques have enabled classification of bio-oil constituents based on functional groups, such as carboxylic acids, sugars, alcohols, ketones, aldehydes, phenols, *etc*. However, such studies only provide information on functional groups representative of the bulk bio-oil mixture. Gas chromatography–mass spectrometry (GC-MS) has been used to provide molecular-level information on low molecular mass, volatile components of bio-oils, and

by this approach more than 300 chemical compounds resulting from degradation of lignin (*e.g.*, ethylbenzenes, xylenes, benzenediols, benzaldehydes, *etc*.), as well as cellulose and hemicellulose (*e.g.*, levoglucasan derivatives of furan, glycoldehyde, *etc*.) have been identified.[10,14,18,21–25] A majority of compounds identified by GC-MS have molecular mass < 200 (g/mol). However, analyses involving gel permeation chromatography (GPC) and thermogravimetric (TG) techniques have demonstrated the presence of a substantial fraction of high molecular mass compounds, with molar mass up to 2,000 g/mol.[24,27] Compounds present in the high molecular weight fraction are relatively less volatile and would not be amenable to GC-MS analysis. Therefore, alternative analytical strategies capable of analyzing less-volatile, larger molecules is important for comprehensive characterization of bio-oils.

High-resolution Mass Spectrometry (HRMS) for the Analysis of Pyrolysis Bio-oil

Enormous analytical challenges are associated with the comprehensive analysis of samples, such as crude oil,[28] natural organic matters (NOMs)[29,30] and pyrolysis bio-oils[31,32] that are compositionally very complex. HRMS analysis combined with efficient data processing and visualization approaches is widely used in the analysis of such samples to provide an overview of molecular composition.[33–41] Fourier transform ion cyclotron resonance (FT-ICR)[42,43] and Orbitrap[44,45] mass spectrometry are the primary HRMS techniques used for analysis of such samples. Mass defect[34,35] and van Krevelen [46,39] analyses, as well as contour plots,[38,47] are the common analytical data processing methods that are used to simplify and visualize mass spectra consisting of hundreds to thousands of peaks.

In recent years, HRMS instruments equipped with a variety of ionization sources, such as, electrospray ionization (ESI),[48–50] laser desorption ionization (LDI),[51] atmospheric pressure photoionization (APPI)[52,53] and atmospheric pressure chemical ionization (APCI)[53] sources have been reported to be useful for more comprehensive characterizations of bio-oil. Accurate mass and elemental composition data derived from HRMS analyses have enabled identification of hundreds to thousands of chemical formulas, representing oxygen-rich components in bio-oils. While knowledge of chemical formulas has greatly improved current understanding of bio-oil composition, structural characterization of chemical species in bio-oils is likely to enable more efficient optimization of catalytic upgrading processes. Unfortunately, the complexity of bio-oils limits the utility of tandem mass spectrometry (MS/MS) to elucidate structural information in the absence of a complementary separation.[31,32] To this end, ion mobility – mass spectrometry (IM-MS)[54] could be useful, and the utility of this technique in promoting a more detailed understanding of bio-oil composition needs to be explored. Additionally, comprehensive structural characterization of bio-oils generated from various biomass resources under different pyrolysis conditions could be expected to facilitate an informed selection of feedstocks as well as optimum pyrolysis conditions. Research along these lines could also catalyze the development of upgrading strategies that result in better quality fuels.

The next few sections are included to provide the reader with a brief introduction to various analytical tools and techniques used to complete the research described in the remainder of this dissertation.

Mass Spectrometry (MS)

Mass spectrometry is an analytical technique that is used to measure masses of molecules.[55] MS has emerged as a powerful analytical technique for both identification and structural elucidation of unknown compounds, as well as quantification of known compounds. There is much variation with respect to MS instruments or techniques; however, all share three basic actions: (1) the ionization of molecules in an ion source primarily with the aid of laser or electric potential, or many other energetic processes such as corona discharge, plasma formation, rapid heating, particle bombardment, *etc*. (2) separation of ions by their mass-to-charge ratio (*m/z*) in a mass analyzer under an applied magnetic and/or electric field, and (3) detection of these ions. Figure 1.1 illustrates the major components and processes in a typical mass spectrometry analysis.

Although MS has been utilized in various fields of physics and chemistry since it was introduced about a century ago,[56,57] the invention of "soft" ionization techniques (*i.e.*, matrix assisted laser desorption ionization, MALDI;[58,59] and electrospray ionization, ESI)[60,61]) has exponentially increased the number of MS applications reported in the last three decades. This increase has been primarily due to the suitability of these techniques for ionization of biomolecules. In recent years, MS has been an indispensable qualitative and/or quantitative analytical technique in various fields focused on understanding complex mixtures (*e.g.*, petroleomics,[28,38,62] metabolomics,[63–68] proteomics,[69–71] lignocellulomics[72–74], *etc*.).

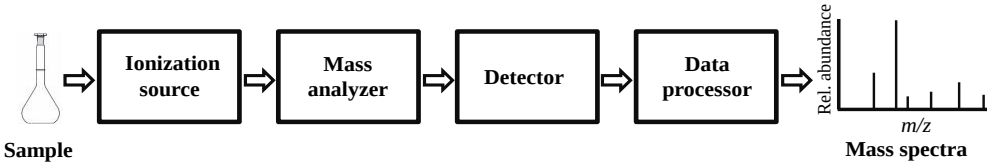


Figure 1.1. Schematic of a mass spectrometer

Electrospray Ionization (ESI)

ESI has been the most widely used ionization technique for the analysis of polar organic and inorganic compounds in mass spectrometry over nearly the past two decades.[75–79] In ESI, pseudo-molecular ions (*e.g.*, deprotonated or protonated molecular ions or non-covalent adducts involving the neutral molecule and some other ion(s), such as, Na^+, K^+, *etc*.) are generated under atmospheric pressure conditions from a sample solution directed through a capillary needle that is maintained at a high electric potential (~1–5 kV).[80] The formation of gas-phase ions from a sample solution occurs in three major steps: (1) production of charged droplets at the capillary tip, (2) evaporation of solvent followed by repeated disintegration of droplets, and (3) formation of isolated gas-phase ions.[80–83]

Three different models have been proposed to describe the mechanism of formation of an isolated gas-phase ions from charged droplets: (i) ion evaporation model (IEM), (ii) charge residual model (CRM), and (iii) charge ejection model (CEM).[83] Schematics of electrospray ionization and three proposed mechanistic models of ion formation are depicted in Figure 1.2. Briefly, according to the ion evaporation model (IEM), an isolated gas-phase ion is generated when a solvated ion leaves the surface of the droplet and solvent is subsequently evaporated. In the charge residual model (CRM), a charged droplet containing an analyte molecule undergoes complete evaporation, leaving behind the charge to the analyte molecule. It is generally believed that ionization of small and relatively polar molecules may follow the IEM mechanism, while relatively large molecules (for example, peptides and proteins) potentially follow the CRM. The charge ejection model (CEM) is a relatively new concept that was introduced to describe

ionization of large amphiphlic compounds. According to this model, ionization occurs by first expelling the nonpolar chain of the molecule from the charged droplet followed by "sequential ejection" of remaining part of the molecule.

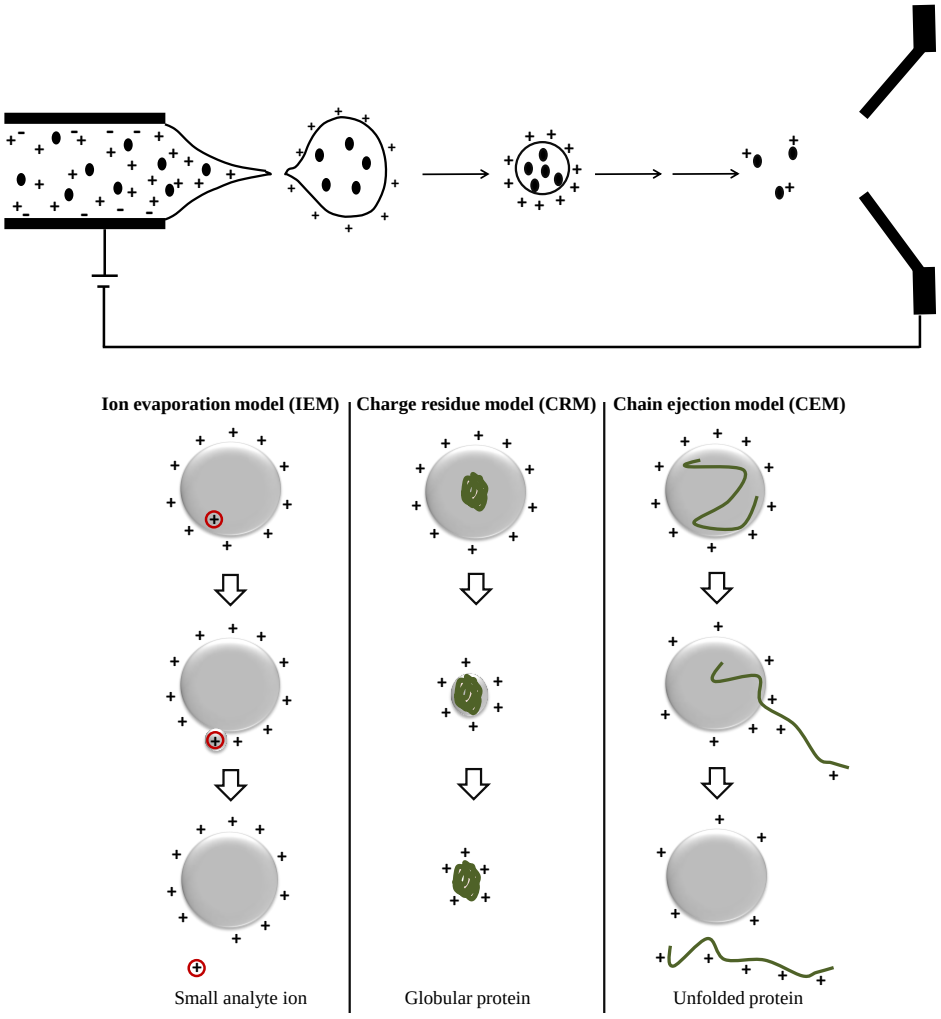


Figure 1.2. Schematic of electrospray ionization (upper panel) and three mechanistic models of ion formation. (Figures modified from reference 83.)

Orbitrap Mass Spectrometry

The first commercial instrument having an Orbitrap mass analyzer was made available by Thermo Scientific (now Fischer Scientific) in 2005.[44] A schematic of a typical Orbitrap mass spectrometer is shown in Figure 1.3.[84] Orbitrap is an advancement

of the Kingdon trap, an ion-trapping device introduced by Kingdon and later improved by Knight.[85] While the Kingdon trap used a thin, straight wire along the axis of a surrounding cylindrical electrode, the Orbitrap mass analyzer consists of three electrodes; a large, spindle-like central electrode and two cup-shaped outer electrodes. The outer electrodes face each other and are separated from each other by a "hair-thin gap" to maintain electric isolation. When voltage is applied between the outer and central electrodes, an electric field is produced that is linear along the axis of the central electrode. The electric field confines ions in the space between the central and outer electrodes and causes them to oscillate harmonically along the central electrode axis. The outer two electrodes work as receiver plates for image current detection.[55]

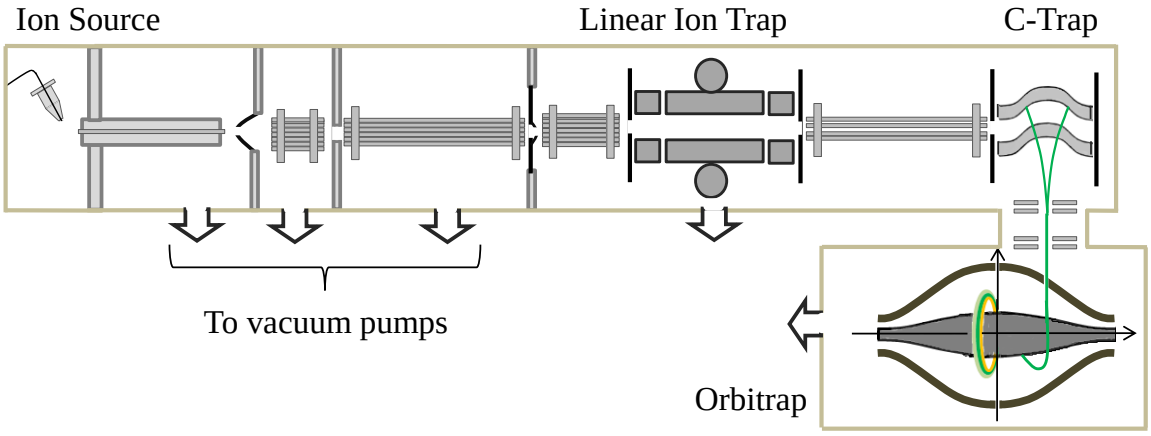


Figure 1.3. Schematic of a linear ion trap-Orbitrap mass spectrometer. (Figure modified from reference 84.)

The overall electric field produced in the Orbitrap is rather complex and termed a quadro-logarithmic field. The resulting potential distribution *U(r, z)* in the mass analyzer is given by the following equation:

$$U(r,z) = \frac{k}{2}\left(z^2 - \frac{r^2}{2}\right) + \frac{k}{2} R_m{}^2 \ln\left(\frac{r}{R_m}\right) + C$$

Where r and z are cylindrical coordinates with z = 0 being the plane of symmetry of the field. The letter k is the field curvature that is determined by the electrodes' shape and applied potential on them, R_m the characteristic radius, and C is a constant. Equation 1 determines the electrostatic field experienced by ions inside the Orbitrap that forces them to move in complex spiral patterns. In this electric field the axial oscillation frequency (ω_z) of an ion is inversely proportional to the ratio of ionic mass to ionic charge (m_i / q) as given in equation:

$$\omega = \sqrt{k \frac{q}{m_i}}$$

Where k is a proportionality constant that depends on the geometric shapes of the electrodes and the applied potential.

Ion detection. The oscillation frequencies of ions are simultaneously determined by detecting the induced image current on the outer electrodes. The time domain signal of the image current is translated to a frequency domain signal by a Fourier transform. Ions of different m/z are separated in the Orbitrap according to their oscillation frequency as the latter depends only on *m/z*.

Time of Flight Mass Spectrometry

The first time of flight (TOF) mass analyzer was introduced by W.E. Stephens in 1946.[55] The basic principles of this analyzer are relatively simple, and mass measurement involves monitoring ions flight times after they are accelerated by a pulsed electric field. When ions are accelerated by an electric field through a field-free region, they all, ideally acquire the same kinetic energy according to the equation:

Electric energy (E_p) = Kinetic energy of ion (E_k)

Or, $qU = ½\ mv^2$

Where, U is the electric field strength and q, m, and v are, respectively, the charge, mass and velocity of an ion. Under this condition, the flight time of an ion (t) is directly proportional to the mass-to-charge ratio:

$$t = k\sqrt{\frac{m}{z}}$$

Where, k is a proportionality constant that depends on instrumental settings and characteristics.

While the configuration of first-generation TOF analyzers was very simple, consisting of a linear drift tube with an electric "pusher" and a detector located at opposite ends, several improvements to this design have been made over the years. For example, most modern TOF analyzers utilize a "reflectron" to achieve better resolving power by narrowing the initial kinetic energy distribution that results, in practice, when ion packets are accelerated by an electric field. TOF MS finds applications in nearly all sub-fields of mass spectrometry. The characteristic features of TOF insturments include (i) a theoretically unlimited *m/z* range and (ii) pulsed analysis, leading to rapid data acquisition. These features are especially important for coupling with MALDI, which is a pulsed ionization technique capable of generating singly-charged ions of large biomolecules.

Collision-induced Dissociation

When mass spectrometry analysis involves, so called, "soft" ionization (*i.e.*, when molecular or pseudo-molecular ions predominant in the ionization step), fragments are

often produced downstream in the instrument by colliding an accelerated ion(s) with neutral gas molecules (*e.g.*, He, N_2, Ar). This approach to fragmentation is called collision-induced dissociation (CID).[55] In this technique, a precursor ion is typically mass selected and fragmented in a collision cell to produce a spectrum of fragment ions. Controlled CID experiments produce fragment ions which are generally characteristic of a single compound or a group of similar compounds. Therefore, CID experiments are useful for gaining structural information, such as, functional group identification or molecular identification through fragment fingerprinting. In recent years, CID has become an increasingly popular structural analysis technique.[86–91]

High-resolution Mass Spectrometry (HRMS)

Mass resolution is a term used to express an instrument's ability to distinguish between two similar masses. The ability of an instrument to resolve *m/z* peaks is called resolving power. In practice, resolving power is typically calculated from the width of a single peak at half height, according to the following equation:

$$R = \frac{m}{\Delta m_{50\%}}$$

Where, $\Delta m_{50\%}$ is the full width of the peak at half maximum.

High resolving power is important to ensure only one kind of ion contributes to a given measurement. It is often a requirement in the analysis of compositionally complex samples where ions with small *m/z* differences can potentially be present.[92] Although high resolution does not guarantee high mass accuracy, the characteristically narrow peaks in a high-resolution spectra reduce the ambiguity of peak centroids and enable better measurement accuracy when a proper calibration strategy is applied. Thus, with

proper calibration, HRMS enables accurate mass measurements that support assignments of molecular formulas to detected analytes.[93]

Mass Measurement Error

Mass measurement error (or accuracy) is generally expressed in parts per million (ppm), and it is calculated as:

$$Mass\ measurement\ error\ (ppm) = \frac{(m/z)_{obs} - (m/z)_{th}}{(m/z)_{th}}$$

Where, $(m/z)_{obs}$ *and* $(m/z)_{th}$ are, respectively, the observed and theoretical *m/z* (in Daltons) of the assigned chemical formula.[94] Modern instruments now routinely generate *m/z* data with mass measurement errors < 5 ppm. For example, mass measurement errors less than 2 ppm are routinely achieved with Orbitrap analyzers,[95] and errors less than 1 ppm have been reported in many FT-ICR[96] analyses. When a high level of mass measurement accuracy is used in combination with carefully developed rules to select the most likely compounds present in a given sample, it is generally possible to assign a single chemical formula to low mass ions ($m/z \leq 500$) with reasonable confidence.

Commonly Used Data Analysis Tools in High-resolution Mass Spectrometry

Mass Defect Analysis

Kendrick mass defect (KMD) analysis. Kendrick mass defect (KMD) analysis[34] is a commonly used tool to identify patterns of elemental composition in a high-resolution mass spectrum of a compositionally complex sample. Kendrick mass for a given homologous series is a normalized *m/z*, and is calculated by multiplying observed mass by a normalization factor, which is a ratio of nominal mass to measured accurate *m/z*

(*e.g.*, 14/14.01565 for a CH_2-homologous series). The difference between this normalized *m/z* with nominal mass is called KMD.

i.e., Kendrick mass defect = Nominal observed mass – Kendrick mass

Where, for a CH_2 homologous series:

Kendrick mass (CH_2) = Observed *m/z* x 14/14.01565

Theoretically, all members of a given homologous series would be expected to have the same KMD value. However, a deviation of 0.001 Da, relative to the theoretical value, is generally accepted in HRMS analysis.

Higher order mass defect analysis. Higher order Kendrick mass defect (MD^2) analysis was introduced by Roach *et al.* and it enables a better visualization of shared homology between complex samples by clustering all members of a homologous series in a single data point. [35] In this analyis, an initial mass defect transformation is followed by a second transformation using a different base group (*e.g.*, CH_2 followed by O) such that an entire homologous series is represented as a single data point that is defined by two unique elemental features. A visual comparison of data resulting from KMD and MD^2 analysis are given in Figure 1.4.

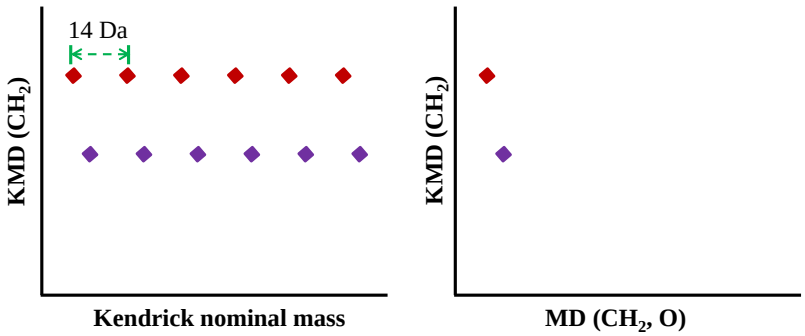


Figure 1.4. Schematic of Kendrick mass defect analysis (left panel) and a second order mass defect analysis (right panel).

Contour Plots

A contour plot is a graphical representation of double-bond equivalents (DBE) *vs* carbon number of identified chemical formulas. DBE represents the number of double bonds plus rings in a given chemical formula and is calculated by the following equation for formula $C_cH_hN_nO_o$:

$$\text{DBE} = C - \frac{h}{2} + \frac{n}{2} + 1$$

DBE values are important to gain preliminary structural information from chemical formulas.[97] For example, a high DBE value means a high degree of unsaturation, which may indicate the presence of a benzene or aromatic ring(s) in the structure of a molecule. In the context of lignocellulosic degradation products, DBE values can be used to correlate products to the polymeric constituents of biomass from which they were derived. For example, a low DBE value may indicate that the resulting product is potentially from cellulose or hemicellulose while a high DBE value may indicate that a product resulted from degradation of lignin.[48]

van Krevelen Diagrams

A van Krevelen diagram classifies compounds based on the distributions of oxygen and hydrogen relative to carbon (*i.e.*, the O/C and H/C ratios). [39,98] Figure 1.5 illustrate the regions in a van Krevelen plot that would be occupied by various compound classes.[94] Historically, van Krevelen diagrams were generated by plotting the elemental O/C and H/C ratio data derived from bulk elemental analysis of a sample. However, it is now common to generate van Krevelen diagrams using HRMS data. Additional information (*e.g.*, the relative abundance of different classes of compounds in a sample)

can be derived from three-dimensional representations of a van Krevelen diagram that include relative ion intensities along the third axis.

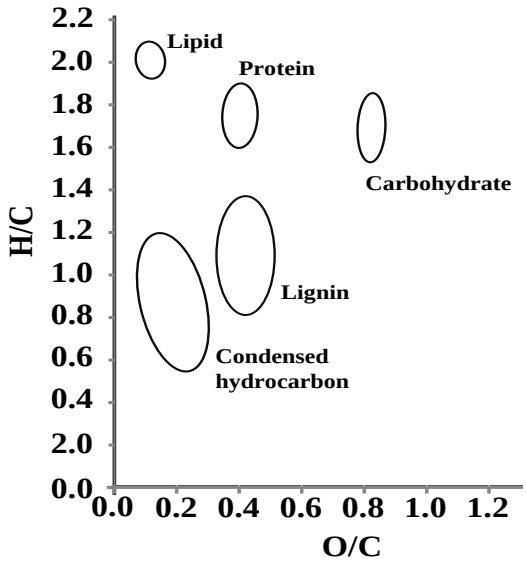


Figure 1.5. van Krevelen diagram showing various regions in H/C versus O/C plot where different compound classes may appear.

Ion Mobility – Mass Spectrometry (IM-MS)

Ion mobility spectrometry (IMS) is an analytical technique that separates gas-phase ions based on shape (or size)-to-charge ratio.[99,100] The millisecond time-scale of IM separations combined with the ability to provide structural information through the shape/size-dependent mobility of ions, makes this technique an attractive complement to traditional separation techniques, such as liquid chromatography or capillary electrophoresis. In addition to providing structural information (*e.g.*, mean cross-sectional area of molecules), the ability of IM-MS to resolve ions that may be indistinguishable by mass spectrometry alone makes IM-MS a powerful tool in the analysis of complex mixtures. Additional structural information may be obtained by investigating mass – mobility correlations.[101] IM-MS has been used to accomplish analytical-scale separations

and structural elucidations in a wide range of fields including petroleomics,[102–104] metabolomics,[105,106] and proteomics[107,108]. There are two basic IMS instrument configurations that are commonly interfaced with commercial mass analyzers: (1) drift tube IMS[99] that emply only DC voltage and (2) travelling wave IMS[104] that utilizes both DC and RF voltages, are commercially available. The latter was used exclusively to complete work described in this dissertation.

A commercially-available IM-MS instrument that utilizes a travelling wave ion guide is shown schematically in Figure 1.6. In this instrument, ions generated at the source are first collected in the trap cell that precedes the mobility region of the instrument, and these ions are introduced in packets to the travelling wave ion guide where the mobility separation occurs. Then, mobility-seprated ions are then transferred to the TOF mass analyzer. The purpose of He-cell situated prior to the mobility cell is to reduce the kinetic energy of ions *via* collisional cooling. Such cooling minimizes ion scattering and/or fragmentation by providing a "softer" entry into the high-pressure IM

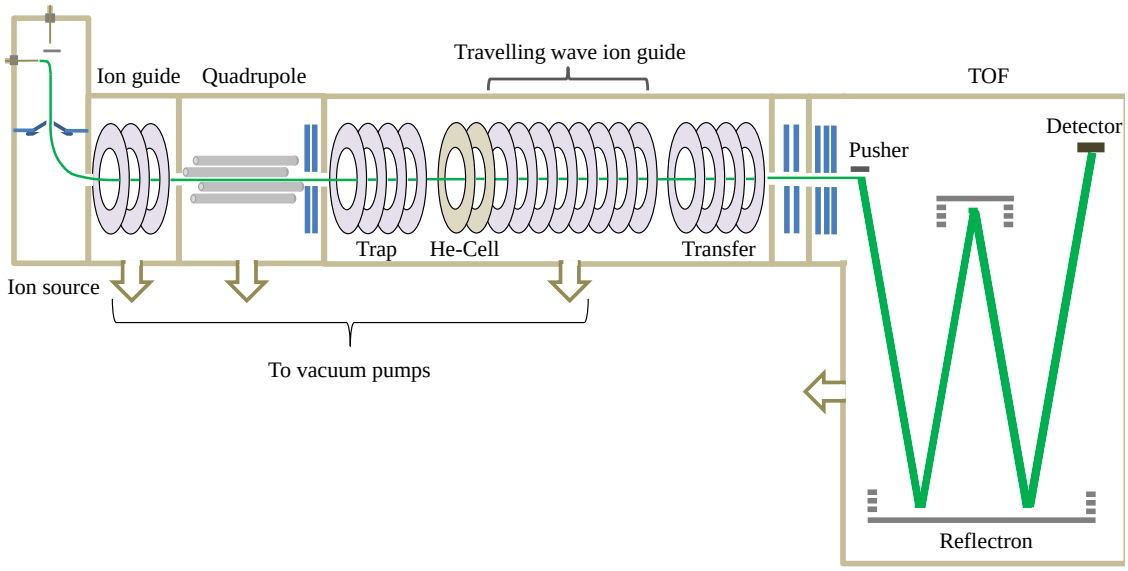


Figure 1.6. Schematic of the Waters Synapt G2-S ion mobility mass spectrometer.

environment. Note that the design of this instrument provides the option of performing CID experiments before or after the mobility separtion (*i.e.*, in the trap or transfer cells). This can be beneficial in attempts to gain additional structural information.

Travelling Wave Ion Mobility Spectrometry(TWIMS)

A travelling wave ion mobility spectrometry (TWIMS) device consists of a sequence of stacked ring-shaped electrodes.[109,110] In this ion guide, opposite phases of radio frequency (RF) voltage applied to adjacent rings radially confines ions within the device. A pulse of electric potential is created by applying direct current (DC) voltage to a pair of adjacent rings. Sequentially stepping this DC voltage to adjacent pairs of rings generates a travelling wave of DC potential (*i.e.*, a T-wave) that pushes ions along the axis of the ion guide. These travelling waves propel gas-phase ions through a "sea" of buffer (or drift) gas, which impedes the forward motion of analyte ions due to collisions. Smaller ions experience relatively few collisions and are carried along by the DC potential wave. Larger ions, on the other hand, experience many collisions, the sum of which can result in the ion "rolling over" the wave it is currently "surfing" on. Ions that roll over a wave must wait for the subsequent DC potential wave before they can again move forward along the ion guide. This process results in a size-based separation that depends, fundamentally, on the molecular collision cross-section of analytes. The average molecular collision cross-section in TWIM is given by the following equation:

$$\Omega = \frac{(18\pi)^{1/2}}{16} + \frac{ze}{(k_b T)^{1/2}} \frac{1}{\mu^{1/2}} \frac{760}{P} \frac{T}{273.2} \frac{1}{N} A t_d^B$$

Where, e is the charge on an electron, z is the charge state of the analyte ion, k_b is the Boltzmann constant, N is the number density of the drift gas, T is the temperature, P is the buffer gas pressure, t_d is the drift time, and μ is the reduced mass of the ion and neutral. A and B are correction factors resulting from the nonlinear and time-varying electric field present in the traveling wave device. A visual depiction of TWIMS separations is shown in Figure 1.7.[111]

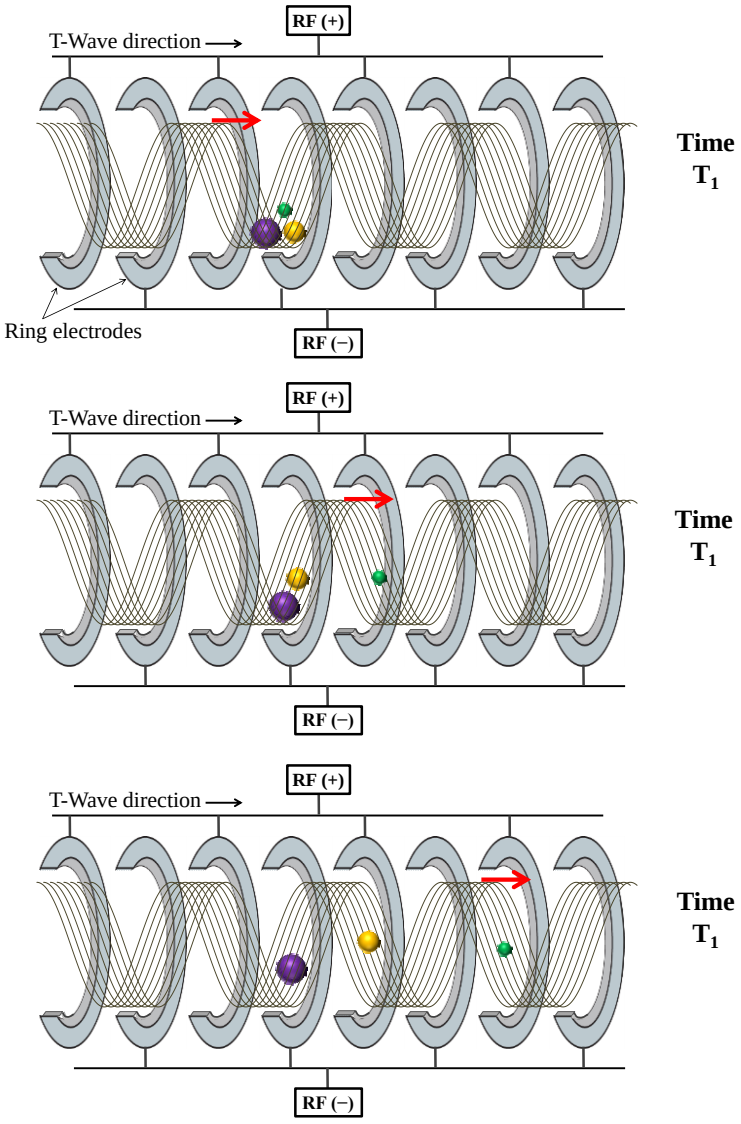


Figure 1.7. Schematic of travelling wave ion mobility spectrometry (TWIMS). (Figure modified from reference 111.)

Mass – Mobility Correlation

IM-MS data is typically visualized in a plot of arrival (or drift) time versus *m/z*. It is common in these plots for peaks to appear along roughly linear trendlines. This is understandable, considering that ion size is generally correlated with mass. Independent trendlines are typically observed for ions of different charge.[112,113] Different compound classes may also appear in unique trendlines.[114] In carefully controlled experiments, trendlines can even serve as a calibration curve to predict the molecular collision cross-section of "unknown" analytes.[115] A representative two-dimensional IM-MS spectra observed in the analysis of sorghum hydrolysate is shown in Figure 1.8.[112] In this dissertation, mass – mobility trendlines are relied on to understand molecular structure as it relates to various homologous series.

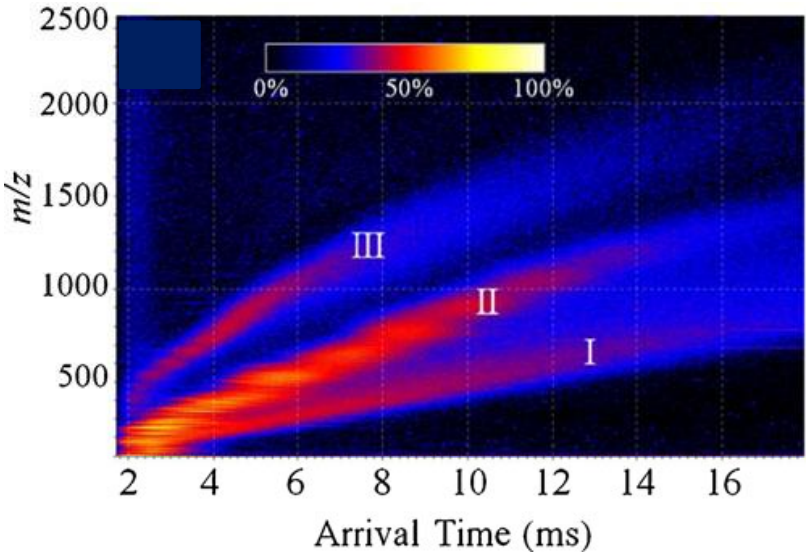


Figure 1.8. Two-dimensional negative-ion mode ESI-IM-MS spectra of $H_2SO_4^-$ pretreated sorghum hydrolysate showing three distinct trendlines resulting from different charge states and compound classes (Figure is reproduced from reference number 112.)

Scope of the Dissertation

The primary contribution of work described in this dissertation is in advancing ion mobility – mass spectrometry (IM-MS) methods for the analysis of compositionally complex mixtures. Pyrolysis bio-oil is used repeatedly as an example of a compositionally complex mixture. Thus, many of the advances developed using this sample type may also be useful in the context of other similar mixtures (*e.g.*, crude oils, natural organic mixtures or hydrolysis products of biomass).

In Chapter Two, IM-MS, high-resolution mass spectrometry (HRMS) and collision-induced dissociation (CID) techniques are applied to characterize molecular constituents of bio-oil. Although HRMS is routinely applied to identify molecular formulas in complex mixtures, this study demonstrates that additional structural information may be obtained from observed mass – mobility correlations and collision-induced dissociation (CID) experiments performed on mobility-separated ions. Chapter two was published as: Dhungana, B.; Becker, C.; Zekavat, B.; Solouki, T.; Hockaday, W. C.; Chambliss, C. K. Characterization of Slow-pyrolysis Bio-oils by High-resolution Mass Spectrometry and Ion Mobility Spectrometry *Energy Fuels* **2015**, *29*, 744– 753. While I performed all the necessary data collection and preliminary analyses and the preparation of initial draft, bio-oil samples were prepared by Dr. William Hockaday. Other co-authors including my advisor Dr. C. Kevin Chambliss mainly contributed in designing experiments and providing feedback on data interpretation as well as manuscript writing. The early response to this approach from the scientific community has been encouraging. For example, a recent review[116] published in *Analyst* states that these experiments represent “technological inroads for IM-MS methods”.

Chapter Three describes fundamental research aimed at understanding the influence of repeat unit and terminal group structure on mass – mobility correlations resulting from analyses of homologous series. Although a few previous studies have highlighted potential analytical applications of mass – mobility correlations in the context of homologous series, a systematic study focused on understanding the influence of chemical structure in such correlations had not been previously reported. Work described in this dissertation suggests that mass – mobility correlations for homologous series could be further developed to facilitate rapid structural profiling in complex mixtures. This chapter has been published online as: Dhungana, B; Becker, C.; Chambliss, C. K. The Influence of Terminal Group and Repeat Unit Structure on Mass – Mobility Correlations Observed for Homologous Series *Int. J. Ion Mob.* December 2015, 1-8. DOI: 10.1007/s12127-015-0187-7. While I performed all the necessary data collection and preliminary analyses and the preparation of initial draft, other co-authors Dr. C. Kevin Chambliss and Christopher Becker mainly contributed in designing experiments and providing feedback on data interpretation as well as manuscript writing.

Sequential solvent fractionation of a bio-oil using traditional techniques and subsequent HRMS and IM-MS analyses are described in Chapter Four. Potential benefits of this analytical workflow include a more comprehensive compositional analysis of bio-oil. This study represents the first application of IM-MS in combination with HRMS for the analysis of solvent fractions of a bio-oil. Overall conclusions of the dissertation and potential areas for future research are discussed in the Chapter Five.

CHAPTER TWO

Characterization of Slow-pyrolysis Bio-oils by High-resolution Mass Spectrometry and Ion Mobility Spectrometry

This chapter published as: Dhungana, B.; Becker, C.; Zekavat, B.; Solouki, T.; Hockaday, W. C.; Chambliss, C. K. Characterization of Slow-pyrolysis Bio-oils by High-resolution Mass Spectrometry and Ion Mobility Spectrometry *Energy Fuels* **2015**, *29*, 744– 753

Abstract

Bio-oils produced from biomass pyrolysis are an attractive fuel source that requires significant upgrading. Before upgrade strategies can be developed, the molecular composition of bio-oils needs to be better understood. In this work, oily and aqueous fractions of bio-oils produced by slow pyrolysis of two feedstocks, pine shavings (PS) and corn stover (CS), were analyzed by negative electrospray ionization (ESI)-Orbitrap and ion mobility-time-of-flight mass spectrometry (IM-TOF-MS). Analyte ion signal was observed primarily between *m/z* 80 and 450 in the mass spectra of these samples. Mass defect analysis and collision-induced dissociation (CID) experiments performed on mobility-separated ions indicated a high degree of homology among bio-oil samples produced from both feedstocks. Oxygen-rich species having between 1 and 9 oxygen atoms and with double bond equivalents (DBEs) ranging from 1 to 15 were identified, indicating that catalytic upgrading will likely be required if slow-pyrolysis bio-oils are to be utilized as fuel. IM-MS and IM-MS/MS analysis of ions belonging to select CH_2-homologous series suggest that mass – mobility correlations and post-ion mobility CID

mass spectra may be useful in defining structural relationships among members of a given Kendrick mass defect series.

Introduction

Pyrolysis converts biomass into three types of potentially useful products: solid bio-char, liquid bio-oil, and combustible gases (*e.g.*, synthesis gas).[117] Pyrolysis of biomass can be classified into two major types: fast (or flash) and slow pyrolysis. Fast pyrolysis involves heating biomass on a time-scale of seconds to temperatures ranging from ~400 to 900 ^{o}C. In contrast, slow pyrolysis typically involves a slow heating rate (< 10 ^{o}C/min) and heat treatment times ranging from a few minutes to a few days. Pyrolysis conditions significantly affect the product distribution. For example, the proportion of bio-oil produced during fast pyrolysis (up to 75% by dry weight of biomass) is substantially higher than that resulting from slow pyrolysis (typically ≤ 30% by dry weight of biomass).[11]

Bio-oils can be catalytically upgraded to useful fuel or value-added fine chemicals, and a number of upgrading techniques have been reported in the literature.[7,118,119] A current aim in bio-oil research is to develop suitable techniques of pyrolysis and catalytic upgrading to make the production of biofuel cost-effective.[120,121] One of the major hindrances toward achieving this goal is a poor understanding of the composition of bio-oils at the molecular level.[31] A number of recent studies have reported on the composition of pyrolysis bio-oils, and these works have been covered in two recent reviews.[31,122] Considerably less effort has been directed at understanding the composition of bio-oils resulting from slow pyrolysis, likely due to inferior bio-oil yield per weight of biomass. However, depending on how catalytic upgrading of a bio-oil is

affected by composition, bio-oil yield per weight of biomass may not correlate directly with downstream fuel content or net value of the product stream. For example, water content of fast-pyrolysis bio-oil is generally high compared to slow-pyrolysis bio-oil,[15] and a slow pyrolysis optimized for synergistic production of bio-char, bio-oil, and synthetic gases could be an attractive (or minimally complementary) alternative to fast pyrolysis for cost-effective biofuel or chemical production.[4]

Information from analyses of bio-oil by conventional techniques is limited to either bulk properties (*e.g.*, functional group identification of the bulk bio-oil by nuclear magnetic resonance[21,123] or Fourier transform infrared spectroscopy[124]) or volatile compounds (by gas chromatography-mass spectrometry[24,125]). Recently, high- and/or ultra-highresolution mass spectrometry (HRMS) coupled to electrospray ionization (ESI),[48–50,126–129] laser desorption ionization (LDI),[51,128] atmospheric pressure photoionization (APPI),[126] and atmospheric pressure chemical ionization (APCI)[53,130] sources has been reported as a useful technique for more comprehensive analysis of bio-oils. Accurate mass and elemental composition data derived from HRMS have enabled identification of hundreds to thousands of chemical formulas, representing oxygen-rich components in bio-oils. While knowledge of chemical formulas has greatly improved current understanding of bio-oil composition, structural characterization of chemical species in bio-oils are likely to enable more efficient optimization of catalytic upgrading processes. Unfortunately, the complexity of bio-oils limits the utility of tandem mass spectrometry (MS/MS) to elucidate structural information in the absence of a complementary separation.[32]

Ion mobility (IM) spectrometry separates gas-phase ions by their shape/size-to-charge ratio, and IM-HRMS instruments are commercially available.[54,110] This combination of techniques reduces compositional complexity prior to mass spectral analysis. For example, nominally isobaric ions with appreciable structural differences can be separated in the mobility dimension. Collision-induced dissociation (CID) experiments can then be performed on mobility-separated ions to inform structure *via* mass spectral analysis of resulting fragment ions. Additional structural information may be obtained from observed mass – mobility correlations. Thus, IM-HRMS is a promising technique for detailed characterization of bio-oil constituents.

In the current work, we present HRMS and IM-MS data obtained from traditional, slow pyrolysis of two feedstock types: pine shavings (PS) and corn stover (CS). Unique elemental compositions were assigned to hundreds of molecules in these samples using standard HRMS techniques. A comparative study of the molecular compositions of oily and aqueous fractions from PS and CS bio-oils was performed with the aim of gaining insight into how bio-oil composition may be correlated with feedstock type. Lastly, the potential of IM-MS and IM-MS/MS to inform structural relationships between adjacent members of a CH_2-homologous series, identified by mass defect analysis, was evaluated. These data represent the first application of IM-MS to bio-oils and provide a valuable point of reference to compare against other pyrolysis conditions.

Experimental Section

Preparation of Slow-pyrolysis Bio-oils

Slow-pyrolysis bio-oils were generated in a custom-built bio-char reactor. In a typical pyrolysis experiment, ~2 kg of biomass was heated in an airtight, 20-L stainless steel reaction vessel. Biomass temperature was monitored and recorded at a frequency of 16 millihertz using a K-type thermocouple placed in the center of the pyrolysis chamber. Corn stover (CS) was heated to a maximum temperature of 500 °C (avg. heating rate ~3.9 °C/min), and pine shavings (PS) were heated to 400 °C (avg. heating rate ~2.5 °C/min). Total reaction time (including the initial temperature ramp) was 300 and 400 minutes, respectively, for CS and PS pyrolysis. Oil condensation was accomplished at ambient air temperature (~18 ^{o}C and ~11 ^{o}C for CS and PS, respectively) in a "cold finger" heat exchanger constructed from 12 linear feet of 316 stainless steel pipes. The oils were inherently biphasic and separation of the phases was accomplished by aspirating the upper (aqueous) layer. All pyrolysis liquids were stored in airtight glass jars at –20 °C until further use.

Data Acquisition

Prior to MS analysis, 50 mg of each bio-oil sample was dissolved in 5 mL methanol, sonicated for 5 seconds and then filtered through a 0.2-µm Acrodisc syringe filter (Pall Corp., Port Washington, New York, USA). Filtrate was further diluted with methanol to a final concentration of 1 mg/mL. During data acquisition, a mass spectrum of methanol diluent was acquired before analysis of each bio-oil sample. The same methanol diluent was analyzed after each bio-oil sample until the system was verified to

be clean by careful comparison of methanol spectra before and after the analysis of each bio-oil sample by the instrument.

Full-scan mass spectra were obtained using a linear ion trap-Orbitrap mass spectrometer (LTQ-Orbitrap Discovery, Thermo Scientific, San Jose, CA) equipped with an ESI source. Consistent with previous analyses of bio-oils, negative-ion mode ESI was utilized in these studies. Negative ESI-HRMS data were acquired by direct infusion at a flow rate of 10 μL/min. ESI and source parameters were carefully optimized to enhance the instrument's sensitivity in the mass range over which we observed peaks from bio-oils (ca. *m/z* 50 to 450). Optimized electrospray source conditions that were used are as follows: sheath and auxiliary gas flow 80 and 5 units, respectively; heated capillary temperature 275 ^{o}C; electrospray voltage –4 kV; transfer capillary voltage –45 V; tube lens –128 V. The Orbitrap mass spectrometer was operated at a mass resolving power of 30,000 (at *m/z* 400; $R = m/\Delta m_{50\%}$) to acquire full scan mass spectra in the *m/z* range of 50 to 500. For each bio-oil sample, data were collected for 200 scans with a maximum injection time of 300 milliseconds. Each bio-oil sample was spiked with sodium dodecyl sulfate (SDS) (final concentration 0.1 ppm) prior to infusion. Mass was locked at the accurate mass of the SDS negative ion, *m/z* 265.1479, for internal mass calibration. With this strategy, high mass accuracy (e.g., average RMS error < 1.5 ppm, Table A.1) was generally possible. Accurate masses of spectral peaks having a signal-to-noise ratio (S/N) greater than 15 were exported from Thermo Xcalibur (Ver. 2.2) to a spreadsheet for subsequent data analysis. Note that the majority of background peaks observed in mass spectra of the methanol diluent were below this S/N threshold. A few background peaks had S/N greater than 15 in the bio-oil spectra; those peaks were identified by comparison

and removed from the spreadsheet before further analysis. Additional benefits of using an S/N threshold are twofold: (1) the most abundant peaks are expected to better represent the majority of the bio-oil mass (by weight), and (2) we believe these peaks will be most important towards initial valuations of fuel quality/potential.

IM-MS and all CID data were obtained on a Synapt G2-S HDMS instrument (Waters Corp., Manchester, UK) equipped with a nano-ESI source operated in negative ionization mode. The trap and transfer cells were maintained at 4.0 and 2.0 V, respectively during IM-MS analysis. A travelling wave height of 40 V and wave velocity of 1,000 m/s was applied to the mobility drift cell for all analyses except CID experiments. Other parameters used during IM-MS and IM-MS/MS are given in Table 2.1. IM-MS/MS experiments were conducted by mass selecting the nominal mass of interest, followed by IM separation and subsequent fragmentation in the transfer cell. Mass accuracy for all identified fragment ions was under 10 ppm error, consistent with results typical of TOF CID spectra. During IM-MS/MS experiments, the transfer cell and wave velocity were set to 40 V and 1200 m/s, respectively, while the wave height was individually adjusted to provide optimal separation and peak shape of precursor ions. IM-MS data were acquired with the TOF analyzer operated in high-resolution mode ($m/\Delta m_{50\%}$ ~40,000 at *m/z* 300). IM-MS data were processed using MassLynx software (Ver. 4.1).

Mass Defect Analysis and Assignment of Chemical Formulas

Kendrick mass defect (KMD) analysis[28,34] was performed to identify molecular series that differed by the accurate mass of a specific functional group (*e.g.*, CH_2, O, *etc.*) and to confidently assign molecular formula to higher *m/z* values (*i.e.*, where several

chemical formulas were possible). In general, Xcalibur identified a unique molecular formula for the lowest *m/z* member of a given molecular series when the following assumptions were applied: (1) observed *m/z* values correspond to ions with an even number of electrons,[132] (2) C ≤ 40, H ≤ 80, O ≤ 15, N ≤ 5, and (3) mass accuracy ≤ 3 ppm error. A molecular formula was assigned to every other member of the same series by adding one CH_2 unit per 14 Da increase in observed nominal mass. The double bond equivalent (DBE) value for molecular formula $C_cH_hO_oN_n$, representing the number of double bonds + rings, was calculated as described previously.[31]

Table 2.1. ESI and IM parameters of Synapt G2-S used in this study

Instrument Parameter	Value
ESI Infusion Rate (μL/min)	0.5
ESI Voltage (kV)	2.7
Source Temperature (°C)	90
Desolvation Temperature (°C)	150
Sampling Cone Voltage (V)	30
Extraction Cone Voltage (V)	35
IM Cell Pressure (mbar)	3.1
Helium Cell Pressure (mbar)	1.4×10^{3}
Trap Cell Pressure (mbar)	2.3×10^{-2}
Transfer Cell Pressure (mbar)	2.4×10^{-2}

Chemical formulas were not assigned to even mass ($[M - H]^{\cdot -}$) ions in this study. Even mass ion signals were mostly present at low relative abundance, and a majority of these peaks were readily identified as $^{13}C_1$ isotopic peaks (*e.g.*, 209 of 311 even mass peaks in the oily fraction of PS). A few even mass peaks were identified as fragment ions, and many of the remaining peaks were not well resolved with clear shouldering or overlap with $^{13}C_1$ isotopic peaks, especially at higher *m/z* (> 250), leading to mass

errors > 20 ppm for chemical formula assignments. The second order mass defect approach used in this manuscript was described in detail by Roach *et. al.*[35]

Results and Discussion

Molecular Weight Distribution and Compositional Analysis of Slow-pyrolysis Bio-oils

Representative high-resolution mass spectra of the four bio-oil fractions examined in this study (*i.e.*, oily and aqueous fractions from PS and CS) are shown in Figure 2.1. In general, a greater number of spectral peaks (S/N ≥ 15) were observed in the oily fraction of bio-oils from both feedstocks than in the corresponding aqueous fractions. Only singly-charged ions were identified in these spectra, and signal for presumed $[M - H]^-$ ions was observed primarily between *m/z* 80 and 450. We observed a similar molecular weight distribution by ESI-TOF-MS analysis (Appendix Figure A.1), and attempts to optimize the instrumental conditions for higher mass ions did not produce significantly different spectra. Visual comparison of spectra in Figure 2.1 revealed that the overall distribution of ion signals, as well as, the composition of more abundant ions was similar between PS and CS oily fractions. In contrast, noticeable differences in relative abundances were observed between the two aqueous fractions. The latter observation suggests that the composition of water-soluble pyrolysis products may depend on the initial composition of the feedstock. However, it is important to note that a large number of peaks with the same exact *m/z* were present in mass spectra of all four samples. Compositional similarities between the oily and aqueous fractions derived from a common feedstock are not unexpected since polar organic molecules undoubtedly partition between immiscible oily and aqueous phases.

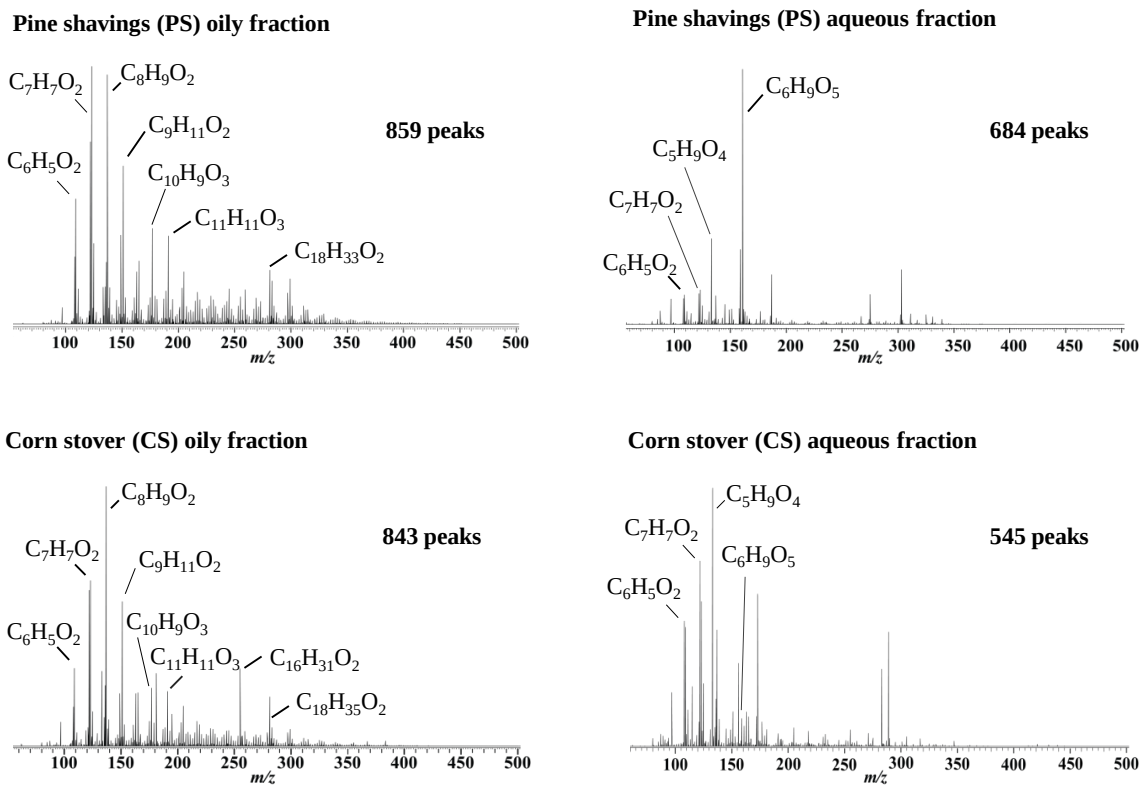


Figure 2.1. Molecular weight distributions observed in (–)ESI-LTQ-Orbitrap mass spectra of bio-oils. Number of analyte peaks with S/N ≥ 15 present in each spectrum is provided at the top right of each spectra.

Figure 2.2 (Panels A and B) shows Kendrick mass defect plots (KMD; plotted as a Kendrick mass of CH_2) for the bio-oil fractions analyzed in the present study. Plots in Figure 2.2 are overlaid to visually compare inter- and intra-sample homologies. The results from KMD analyses indicate the presence of numerous homologous series that are common to both the oily and aqueous fractions. However, the number of analytes in a given series extends to higher *m/z* in the oily fractions, which is consistent with longer aliphatic chains in these sample types. As described in the Experimental section, elemental composition of bio-oil constituents was identified by either accurate mass or a combination of accurate mass and propagation of chemical formula through a homologous series. A maximum KMD spread of 0.001 Da was selected as the criterion

for any two peaks to be considered in the same series, and intra-series KMD values typically varied by less than 0.0006 Da, even when the series contained 10 or more members.

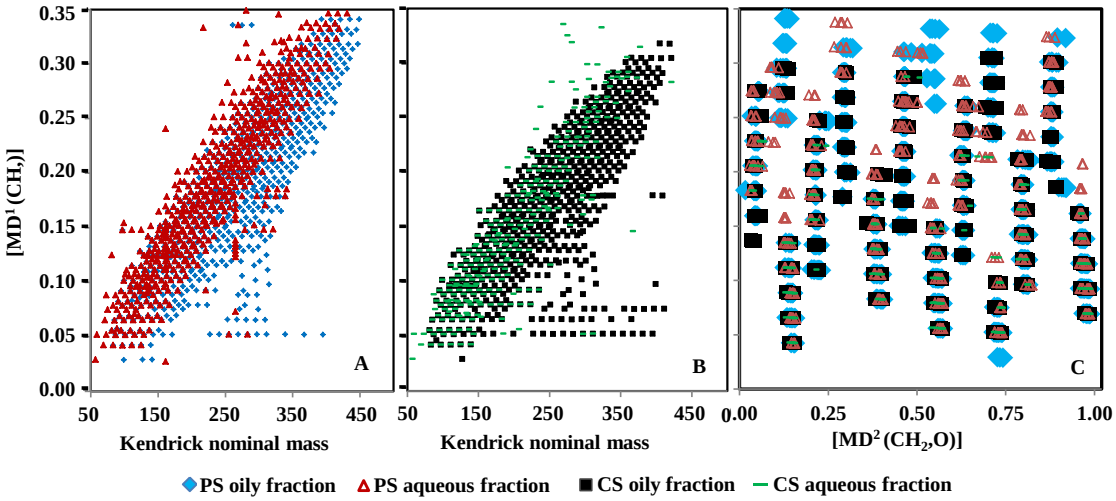


Figure 2.2. First-order mass defect plots of oily and aqueous fractions of bio-oils from pine shavings (PS) (Panel A) and corn stover (CS) (Panel B). Intra-series homology is demonstrated with a second-order mass defect (MD2 (CH_2, O)) plot (Panel C) for all series containing ≥ 3 members.

Generally speaking, these data sets are visually similar to those reported by Smith et al. in their comparative HRMS analysis of fast pyrolysis bio-oils from a red oak feedstock, both in terms of molecular weight distribution and complexity.[48] However, other studies employing negative ESI have reported even higher molecular weight distributions (ca. 150 – 650) and increased complexity relative to the current work.[49,50] To some extent, these differences are a result of the study-imposed peak-selection threshold. It is also possible that these differences arise from the different thermochemical conditions used during the pyrolysis process. For example, relatively slower heating rates (< 4 ^{o}C/min), lower temperatures, and longer vapor-residence times

– all of which were employed in the current study - have been shown to result in lower molecular weight oils.[10,21] In total, we assigned chemical compositions to 735 analytes in the PS oily fraction, 448 analytes in the PS aqueous fraction, 640 analytes in the CS oily fraction, and 239 analytes in the CS aqueous fraction (Table A.2). Carbon, hydrogen, and oxygen atoms were the primary elements observed in assigned molecular formulas, and compounds with two oxygen atoms per molecule were generally more abundant than other O_x classes with respect to spectral intensity (Figure 2.1). A few nitrogen-containing compounds were also observed (*i.e.*, O_xN_y classes of compounds, where, x = 1–4 and y = 2 or 4). These nitrogen-containing species may be of protein origin. However, it is not possible to speculate on their structure further without substantial structural work (*e.g.*, CID, NMR, *etc.*), which is beyond the scope of this paper. We note that they may warrant additional evaluation in the future if they are consistently observed in bio-oil samples. Although unresolved sample constituents can complicate assignments of chemical compositions, we are reasonably confident in the accuracy of molecular formulas reported in this work. Careful inspection of mass spectra between *m/z* 150 and 250 (where $m/\Delta m_{50\%}$ ~50,000) revealed that a mass difference of 0.015 Da, (*i.e.*, C_4 vs O_3) was the smallest interval observed regularly between adjacent peaks (*e.g.*, see Appendix, Figure A.2, Panel A). Two peaks with this mass difference could also typically be identified at higher *m/z* (*e.g.*, Figure A.2, Panel B). In a few instances, the lower mass peak appeared as a low-intensity, unresolved shoulder on a larger peak (*e.g.*, Figure A.3). In this latter case, composition was only assigned to the more abundant component.

To further evaluate the potential of unresolved peaks to influence the accuracy of chemical formula assignments, the PS oily fraction was analyzed on a 9.4 tesla FT-ICR

mass spectrometer capable of mass analysis above *m/z* 300.[133] These ultra-HRMS ($m/\Delta m_{50\%}$ ~150,000 at ~*m/z* 350) data showed that in some cases, a single high-mass peak in Orbitrap mass spectra represented more than one ion. Even then, however, the approach used to assign molecular formulas accurately identified the elemental composition of the more abundant component in the unresolved peak (*e.g.,* Figure A.4). These results (described above) essentially mirror observations by Smith and coworkers, who compared the same model Orbitrap that we use with data collected on a 7 tesla FTICR. For their sample, they noted all peaks were resolved in both instruments when measuring *m/z* 300 or less. Although a few unresolved peaks were detected above *m/z* 300 in the orbitrap data, the authors stated that they were typically very low intensity (< 0.1% relative abundance) and did not affect the "overall picture" of the Orbitrap data. None-the-less, the analyst should always carefully evaluate their sample to ensure data quality. Based on our observations and those described in Ref. 17, it is reasonable to conclude that improved-resolving-power MS could be required when interested in high-mass (*m/z* > 300) and/or low-level analytes in bio-oil samples.

Second order Kendrick Mass Defect (MD^2) analysis was introduced by Laskin and co-workers,[35] providing a clear visual comparison of shared homology between complex samples. In MD^2 treatments, an initial mass defect transformation is followed by a second transformation using a different base group. Figure 2.2C shows the results of MD^2 analysis, using base groups of CH_2 and O, for the bio-oil samples investigated in this study. In this approach, all members of a given homologous series will ideally appear as a single data point or cluster. Modest scatter of data in Figure 2.2C falls within the expected precision of accurate mass measurements used to construct this plot. Majorities

of the observed homologous series were common to all four samples (supporting information, Table A.2). This observation suggests that these bio-oils are similar at the molecular level even though they are derived from different feedstocks. Series belonging only to a single feedstock or to a single fraction (*i.e.*, aqueous or oily) may also be identified in Figure 2.2C. However, the uniqueness of these series is not definitive. Selection criteria for data used in the preparation of Figure 2.2C excluded series containing fewer than three members, as well as ion intensities below the study-imposed signal-to-noise threshold (S/N = 15).

Oxygen Heteroatom Distributions and Isoabundance-Contour Plots of DBE vs. Carbon Number

All chemical formulas identified in the present study contained at least one oxygen atom. Figure 2.3 (Panels A and B) shows the relative abundances of various classes (O_x, x = 1–9) of oxygen-containing compounds observed. Compounds with 2 to 4 oxygen atoms per molecule were most abundant in both oily and aqueous fraction from CS. A slight shift to higher O_x class was observed in the PS aqueous fraction, but overall, the observed heteroatom class distributions were quite similar between the four samples. Although comparison of data obtained using different instruments is not straightforward, a much larger shift to higher O_x class (*i.e.*, from O_4 to O_{10}) was reported by Jarvis *et al.*[49] when comparing oily and aqueous fractions of bio-oil derived from pyrolysis of pine pellets. This observation may support differing compositions of bio-oils produced under different pyrolysis conditions. However, it is also possible that the observed difference is an artifact of differing experimental designs.

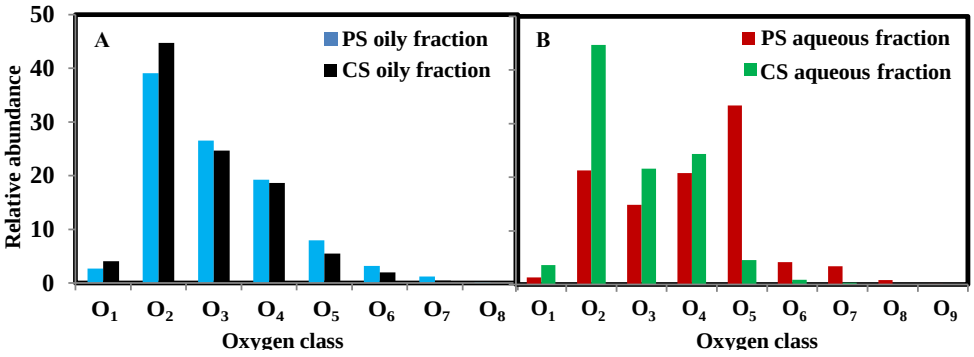


Figure 2.3. Abundance of various O_x compound classes relative to total number of oxygen containing compounds identified in oily and aqueous fractions (Panels A and B, respectively) of pine shavings (PS) and corn stover (CS) bio-oils.

Plots of DBE versus carbon number for compounds containing 1 to 6 oxygen atoms, shown in Figure 2.4, are in some ways qualitatively similar to those reported in previous studies of pyrolysis bio-oils derived from various feedstocks.[48–50] In particular, the general trend of analytes shows an oblong distribution approaching higher DBE values at higher carbon number. Although some ions in both PS and CS oily fractions

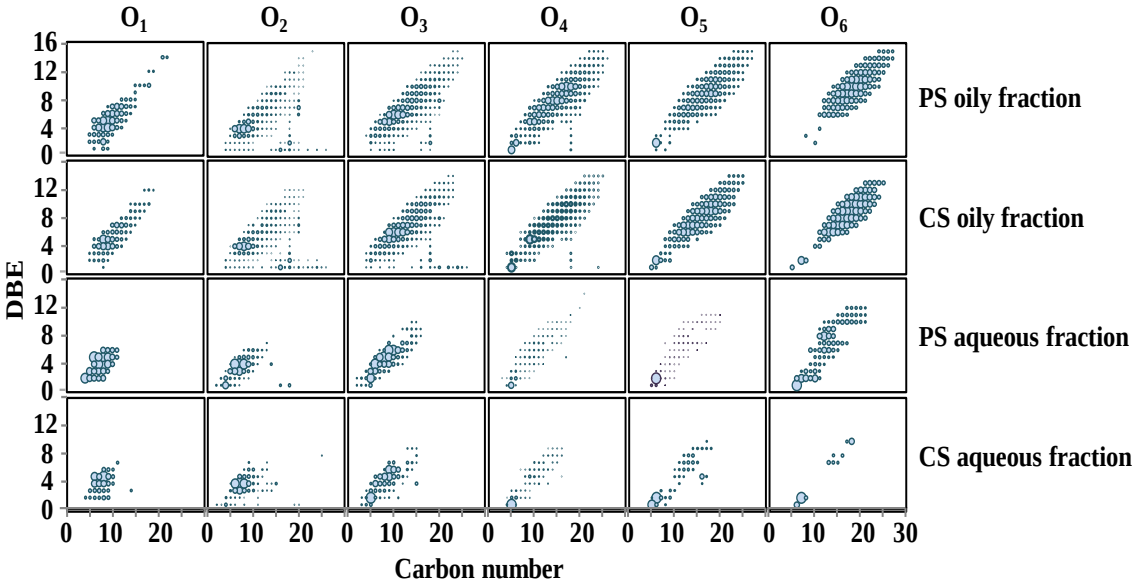


Figure 2.4. Contour plots for heteroatom compound classes O_1 to O_6 in each bio-oil fraction. Bubble size represents the intensity of the corresponding peak in the mass spectrum relative to the most intense peaks in the same O_x class within a given spectrum.

exhibit low DBE (2 or less) at high carbon number (10 or higher), such as, in O_2 to O_4 heteroatom classes, a general increase in DBE with increasing carbon number was observed. In fuel-related samples, DBE and oxygen content are frequently relied on to provide an estimate of fuel quality. For example, DBE represents the degree of unsaturation of a molecule, and higher unsaturation is correlated with decreasing energy content of a potential fuel. In a similar manner, low oxygen content represents better fuel quality. The prevalence of compounds with a high degree of unsaturation and high oxygen content in Figure 2.4 demonstrates that significant upgrading of slow-pyrolysis bio-oils is likely to be required if these liquids are to be useful as fuel.

Some structural information may be inferred from data in Figure 2.4, but it is important to recognize that conclusions drawn solely from HRMS data are somewhat speculative without additional confirmatory analyses. As noted in an earlier study,[49] the general increase in DBE with increasing carbon number is consistent with a polymeric structural motif. Similar data have been used to further classify analytes as having either cellulosic or lignin origin.[48] Such classification is based on the fact that lignin decomposition products are expected to have DBE ≥ 4, while compounds derived from carbohydrate polymers (*i.e.*, cellulose or hemicellulose) are more likely to have DBE < 4 and higher oxygen content per molecule. Data in Figure 2.4 suggest that a majority of compounds in slow-pyrolysis bio-oils that are amenable to ESI are structurally more related to lignin than cellulose. This classification is based on the core structures of lignin and cellulose. However, DBE alone is not a conclusive means of correlating pyrolysis products with the macromolecules of biomass. For example, it is probable that compounds with high DBE may also be formed from cellulosic compounds that undergo

extensive dehydration during pyrolysis. Compounds exhibiting high carbon number and DBE = 1 potentially represent fatty acids.[24] Many analytes in the O_1 class had chemical formulas that were consistent with derivatives of phenol, naphthol or similar small molecules that have been previously observed in bio-oils.[31] It is possible that some analytes in the O_1 class represent fragments of larger molecules, since exhaustive evaluation of in-source fragmentation was beyond the scope of this study and the abundance of O_1 species was generally low relative to ions belonging to higher order oxygen classes.

Structural Insight through IM-MS Mass – Mobility Correlations

The oily and aqueous fractions of bio-oils from PS and CS were analyzed by IM-MS. Representative two-dimensional IM-MS spectra of PS oily fraction is displayed in Figure 2.5A and shows that the majority of ions from PS oily fraction fall along a broad, roughly linear distribution. Similar distributions of ions were also observed in two-dimensional spectra of other bio-oil fractions (Figure A.5). This single distribution is typical when a majority of the sample ions fall within a similar compound class or type (*e.g.*, compounds such as aliphatic and/or aromatic acids, ketones and phenols). Note that the few lower mobility ions (green box in Figure 2.5A) that are separated from the main distribution were also observed in spectra of the methanol diluent. We further interrogated data for PS oily fraction by plotting the centroid of arrival-time distributions of twenty-seven of the most intense CH_2-homologous series in Figure 2.5B (*i.e.*, compounds in the top row of Figure 2.4, belonging to oxygen classes ranging from O_2 to O_5 and having DBE between 3 and 11). Although there is significant overlap between homologous-series distributions of each O_x class, a correlation between class number and

arrival time exists. Specifically, when comparing the arrival times between oxygen classes at a given nominal mass, we generally observe a shift to shorter arrival times with respect to increasing O_x class (*i.e.*, the general trend is such that each subsequently higher O_x class in Figure 2.5B is shifted down). This trend is consistent with more condensed or compact structures and/or fewer atoms per molecule with increasing O_x class. Note that data in Figure 2.5B suggests that higher resolution IMS may be useful when screening for specific heteroatom classes.

Individual trendlines representing mass – mobility correlations of CH_2-homologous series are displayed in Figure 2.5C. First, we observe that trendlines in a given O_x class appear at shorter arrival times with respect to increasing DBE, especially below *m/z* 250. That is, IM-MS data show that higher DBE species have drift times consistent with more compact and/or condensed structures in this mass range, potentially enabling relative DBE evaluations when accurate chemical formulas are not available. Moreover, we note that trendlines corresponding to series with higher DBE exhibit steeper slopes than those of lower DBE in the same O_x class, resulting in the convergence of trendlines at higher *m/z*. Although a specific rationale for this change in slope is not clear, it may be related to the fact that smaller structures at a given starting mass (*i.e.*, those having higher DBE)

The arrival times of ions that were common to all four bio-oil fractions and that represent members of three different CH_2-homologous series (identified by green, blue, and red trendlines in Figure 2.6) are plotted in Figure 2.7. A negligible difference in arrival time was observed for ions that were assigned the same chemical formula in

different samples, suggesting that analytes having the same or very similar structures are present in each sample. The highly linear mass – mobility correlations observed for these

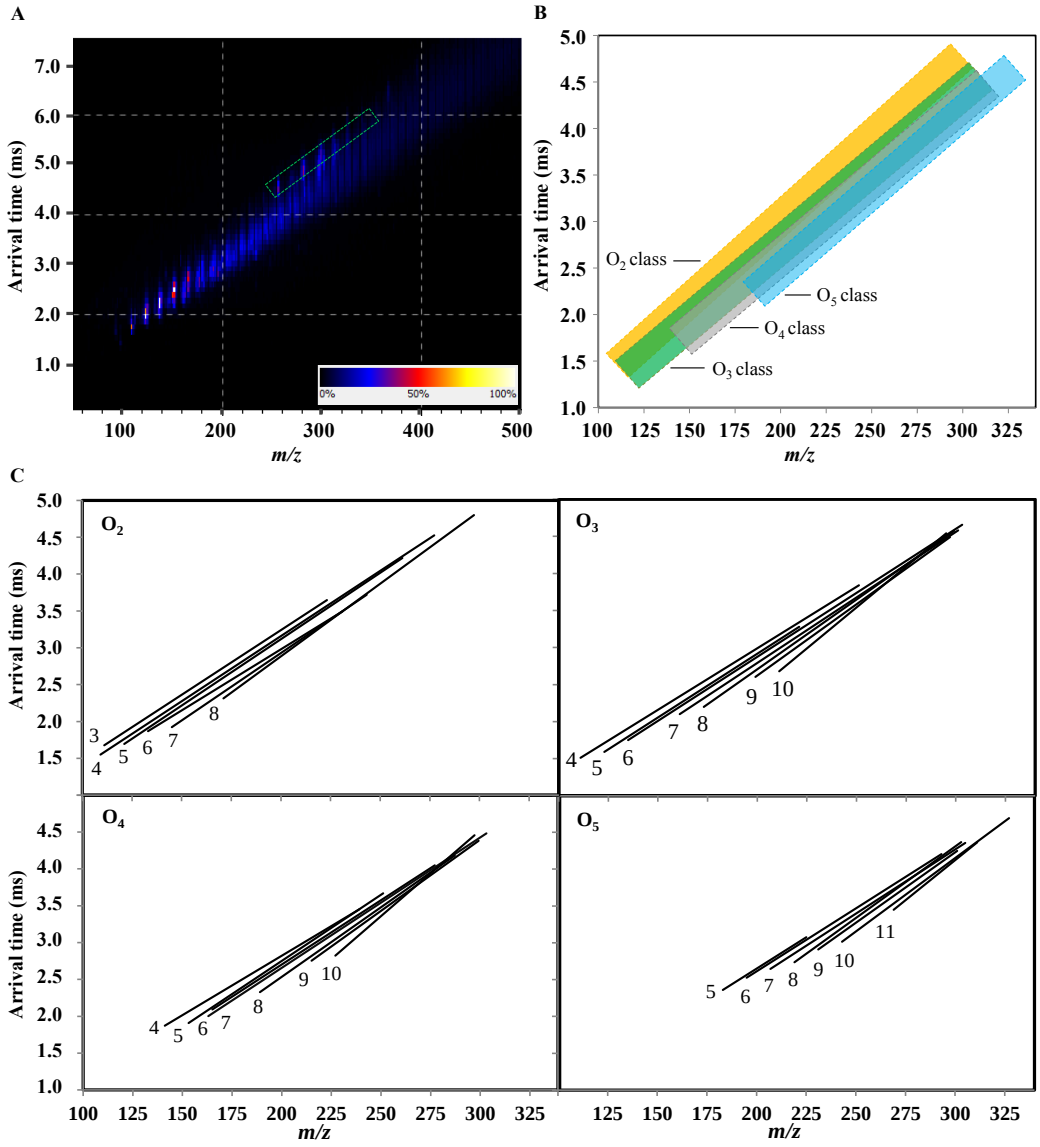


Figure 2.5. (–)ESI-IM-MS data resulting from direct infusion of pine shavings (PS) pyrolysis bio-oil (oily fraction). (A) Two-dimensional IM-MS spectra showing the full MW distribution of the sample. Relative abundance of peaks is designated by coloration. (B) Two-dimensional plot of arrival time versus *m/z* showing the distribution of homologous series within a given oxygen class. Space-filling regions for each oxygen class (O_2 to O_5) were generated by "boxing" arrival-time centroids of analytes observed in the raw spectra. (C) Two-dimensional plots showing the relationship between mass – mobility correlations and DBE within a given oxygen class. DBE of each CH_2-homologous series is indicated by the number at the lower end of each trendline.

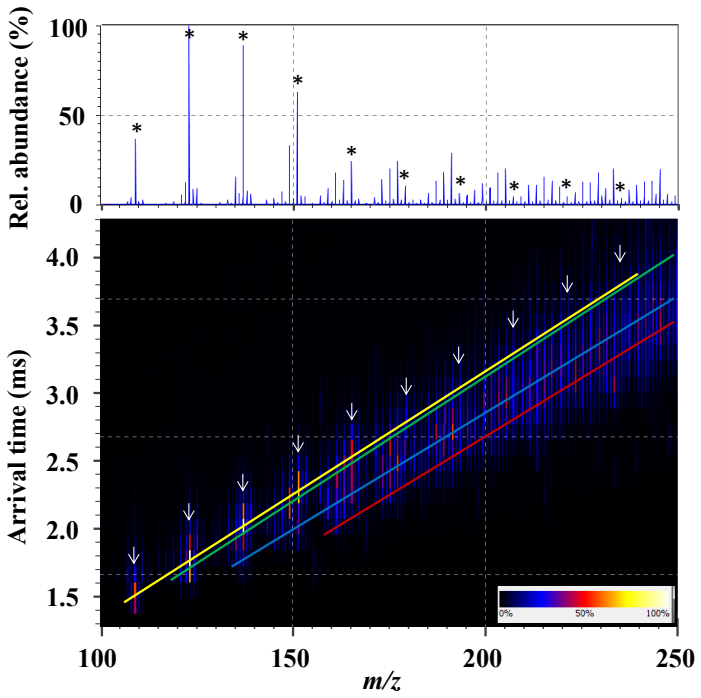


Figure 2.6. (–)ESI-IM-MS data resulting from direct infusion of pine shavings (PS) pyrolysis bio-oil (oily fraction) showing an expanded region between *m/z* 100 and 250 (top panel) and the corresponding two-dimensional IM-MS spectra (bottom panel). The IM-MS spectra are overlaid with trendlines fit to the following CH_2-homologous series: $C_nH_{2n-6}O_2$ (yellow), $C_nH_{2n-8}O_2$, (green) $C_nH_{2n-10}O_3$ (blue), and $C_nH_{2n-12}O_4$ (red). The peaks used to generate the yellow trendline are indicated by arrows and asterisks, respectively, in the top and bottom panels. See text for additional details.

series ($R^2 \geq 0.995$ for all CH_2-homologous series represented in Figures 2.5-2.7), are consistent with structures having a common terminal group (or core unit) and a single, propagating alkyl chain.[115,134] The extent to which IM-MS is able to differentiate between possible isomers in a given homologous series is not fully understood. However, it is reasonable to expect that homologous series of isomers possessing different terminal groups would each follow a unique trendline. This hypothesis is supported by arrival times observed for a commercially-available ketone series (designated by asterisk in Figure 2.7) that represents probable isomers of ions in series 1 (*i.e.*, both series 1 and the commercial ketones share the same chemical formulas, $C_nH_{2n-8}O_2$). The considerable

difference in trendline slopes observed for the ketone series and series 1 provides a remarkable representation of the divergence one could expect if multiple isomers were present at a given *m/*z or if different isomers were present between samples.

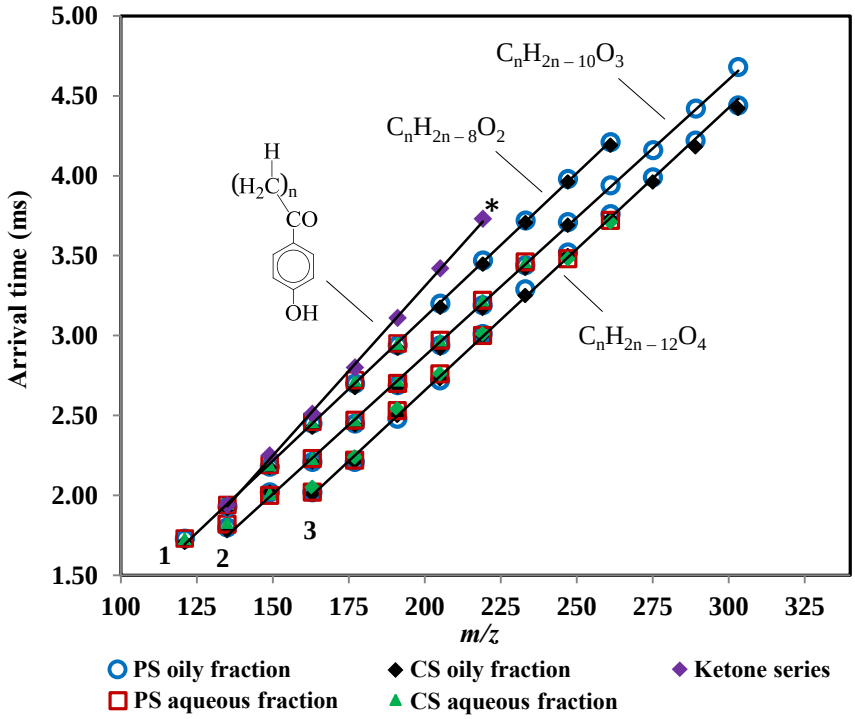


Figure 2.7. Cross-sample comparison of mass – mobility correlations for three CH_2-homologous series present in each of the four bio-oil fractions. The Mass – mobility correlation of a commercially available ketone series (indicated by asterisk) is also shown. Note that this commercial ketone series represents structural isomers of analytes in series 1. Representative IM profiles for series 1 can be viewed in Appendix Figure A.6.

Structural Insight through CID of Mobility-separated Ions

In the current study, CID experiments were performed (1) to assess whether ions assigned the same chemical formula in all four samples were structurally similar and (2) to evaluate the structural relationship between members of a given KMD series. Meaningful interpretation of CID spectra typically requires isolation of a single precursor ion prior to fragmentation. Since multiple peaks were observed at virtually every nominal

mass in the current study (especially above *m/z* 120), we used an IM-MS/MS approach[135] to obtain CID spectra of mobility-separated precursor ions. Figure 2.8 demonstrates the effective separation of two isobaric precursor ions at nominal mass 151 in the mobility dimension. When CID is performed after the mobility separation, arrival time distributions of fragment ions coincide with the distribution observed for the corresponding precursor ion.

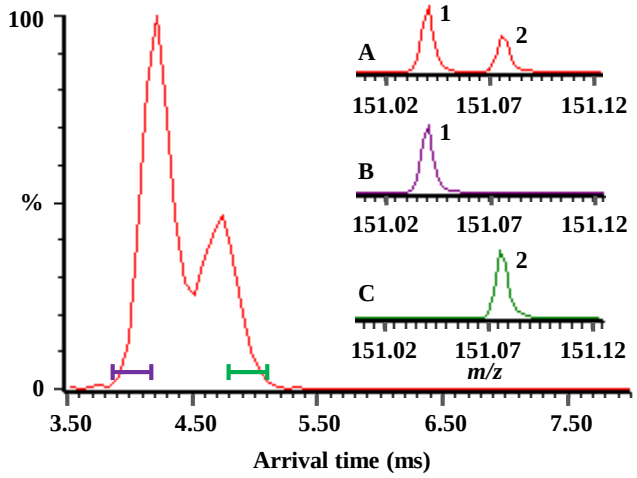


Figure 2.8. IM-MS arrival time distribution of ions at nominal mass 151 in the pine shavings (PS) oily fraction (acquired using IMS wave height = 27 V and travelling wave velocity 1,000 m/s). Inset shows the mass spectrum obtained at nominal mass 151 (Panel A) and the extracted mass spectra from the regions indicated by purple and green lines, respectively, in Panels B and C. Peak 1 was identified as $C_8H_7O_3^-$ (*m/z* = 151.0401) and peak 2 was identified as $C_9H_{11}O_2^-$ (*m/z* = 151.0765).

Figure 2.9 shows representative CID spectra of the two isobaric ions that were separated by ion mobility in Figure 2.8. The observed fragmentation patterns for either $C_8H_7O_3^-$ or $C_9H_{11}O_2^-$ precursor ions (left and right panels, respectively) were nearly identical, independent of sample type. A high degree of similarity was also observed between CID spectra obtained for several other precursor ions that were common to all four samples (data given in Table A.3). Collectively, these data suggest that many, if not

most, molecular formulas assigned to ions that were common to all samples likely correspond to a single structure rather than a distribution of structural isomers. They also serve to reinforce the conclusion drawn from MD^2 analysis above that the bio-oil samples investigated in this work are compositionally similar at the molecular level. Note that integration across small regions of individual peaks does not reveal the presence of additional structures (for example, Figure A.7), which is consistent with our belief that each peak represents a single structure.

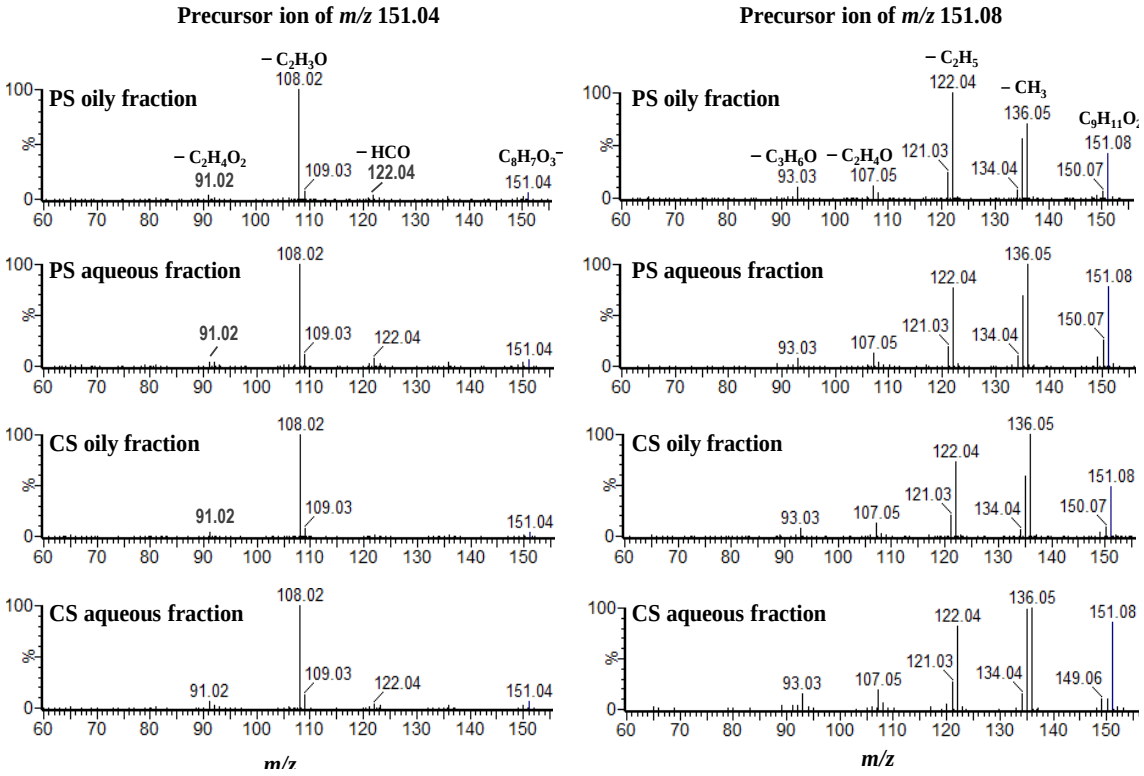


Figure 2.9. CID spectra of mobility-separated precursor ions at m/z 151.0401 and 151.0765 in oily and aqueous fractions of pine shavings (PS) and corn stover (CS) bio-oils.

IM-MS/MS experiments were also performed to further evaluate the structural relationship between members of a given homologous series. Mobility-separated CID spectra are shown in Figure 2.10 (Panels A–E) for the first five members of Series 1 (*i.e.*,

from Figure 2.7). The fragmentation pattern observed for the first member of this series (*m/z* 121.03, Panel B) matched well with spectra we observed for *m*-hydroxybenzaldehyde (*i.e.*, ions with *m/z* 121, 120, 93, 92 and 65, Figure A.8). More significantly, an interesting correlation was observed between increasing number of CH_2 groups and the pattern of neutral loss from precursor ions. In Panel A, the major fragment ions result from neutral loss of HCO and CO, while in Panel B major fragments result from neutral loss of HCO and CH_2CHO. In Panels C-E, the most abundant fragment ions result from overall loss of $C_nH_{2n}CHO$ with n = 1, 2 and 3, respectively. These observations are consistent with the view that members of series 1 have a common terminal unit (*i.e.*, hydroxybenzaldehyde) and a propagating alkyl chain, as illustrated by the tentative structures proposed in Figure 2.10.

Conclusions

High resolution mass spectral analyses of bio-oils produced via slow pyrolysis suggest that these liquids are generally similar in composition to fast-pyrolysis bio-oils. Mass defect analyses, particularly MD^2 analyses, revealed a high degree of compositional similarity in terms of the more abundant components ($S/N \geq 15$) identified in aqueous and oily fractions of CS and PS bio-oils. In contrast, most reports on the composition of fast-pyrolysis bio-oils have emphasized differences between aqueous and oily fractions derived from a common feedstock or between like fractions produced from different feedstocks. In many cases, these differences have been presented with limited consideration of the relative abundance of species being compared. Such a focus may minimize an important reality. Currently, all bio-oils will need to be upgraded if they are

to be used as fuel. Significant advances in catalytic upgrading on a commercial scale are likely to be leveraged by focusing on common constituents that are present at relatively

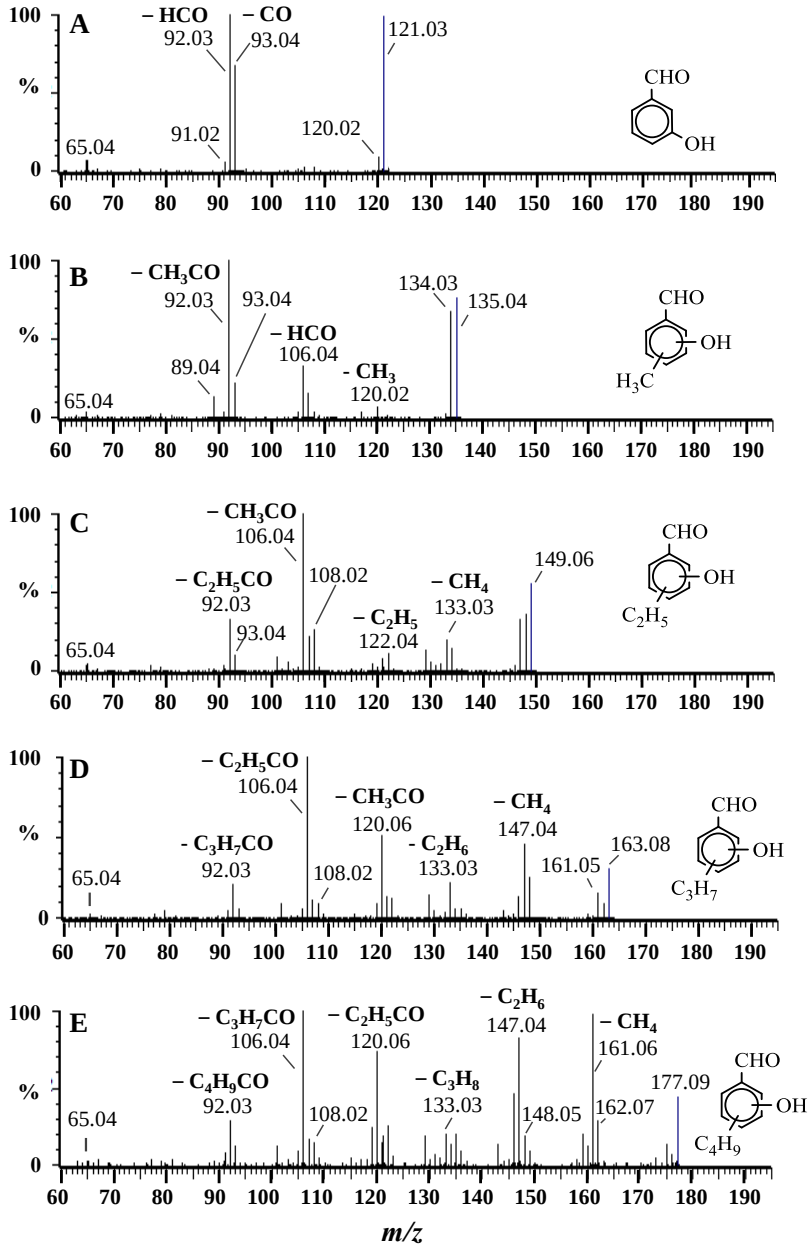


Figure 2.10. Representative CID spectra of the first five members of Series 1 (*i.e.*, shown in Figure 2.7) from the pine shavings (PS) oily fraction. Panels A-E, respectively, represent the 1st through the 5th members of this series. These spectra exhibit similar neutral losses, suggesting structural correlation within this CH_2-homologous series. (Following travelling heights were used to separate precursor ions shown in figure in the mobility dimension: A and B = 25 V, C and D = 28 V, and F = 30 V.)

high concentration. Structural elucidation of the most abundant bio-oil components is also likely to be paramount to the development of optimized catalytic cycles. To this end, IM-MS/MS studies are likely to be a complementary alternative to traditional high-resolution mass spectrometry approaches in the analysis of bio-oils. However, improved resolving power in the mobility and/or mass dimension would be required to achieve comprehensive sample interrogation. This would be especially true for bio-oil mixtures that are more complex than those examined in this work.

Acknowledgments

W.C. Hockaday acknowledges funding from USDA-AFRI-2011-67009-20074. This study was also supported in part by funds from the Faculty Research Investment Program and the Vice Provost for Research at Baylor University. We gratefully acknowledge the Baylor University Mass Spectrometry Facility for continued instrumental support; especially Dr. Alejandro Ramirez for technical assistance. We also thank Dr. Julia Laskin (PNNL) and Dr. Sharon Munisamy (Waters Corp.) for meaningful help concerning data processing and analysis.

CHAPTER THREE

The Influence of Terminal Group and Repeat Unit Structure on Mass – Mobility Correlations Observed for Homologous Series

This chapter published as: Dhungana, B; Becker, C.; Chambliss, C. K. The Influence of Terminal Group and Repeat Unit Structure on Mass – Mobility Correlations Observed for Homologous Series *Int. J. Ion Mob.* December 2015, 1-8. DOI: 10.1007/s12127-015-0187-7

Abstract

Mass – mobility correlations were investigated by ion mobility – mass spectrometry (IM-MS) for 11 structurally-unique homologous series composed of commercially-available compounds. Structural variation involved the inclusion of different repeat units (*i.e.*, glucose, propylene glycol, or ethylene glycol) in oligomeric series and different terminal groups in CH_2-homologous series. Terminal group structures included both aliphatic and aromatic acids, as well as aliphatic and aromatic amines. Mass – mobility correlations were also investigated for select CH_2-homologous series identified in a pyrolysis bio-oil and compared with results observed for commercial series. A linear mass – mobility correlation ($R^2 \geq 0.996$) was established for all series except those in which a substantial change in the gas-phase conformation of ions was probable. Slopes observed for CH_2-homologous series with a single terminal group were significantly steeper than slopes observed for series containing two terminal groups. Additionally, a correlation between slope and double bond equivalents (DBE) suggested that the CH_2-homologous series identified in bio-oil were structurally similar to commercial series containing two terminal groups.

Introduction

Substances derived from natural resources, such as petroleum crude oils or pyrolysis bio-oils, are comprised of a large number of homologous series.[28,136] It is well known that chemical and physical properties (*e.g.*, chemical reactivity, boiling point, solubility, *etc.*) within a homologous series follow general trends. These trends are likely derived from a common structure that is systematically changed by a propagating group of atoms (*e.g.*, CH_2). It is common practice to consider how a structural change is likely to alter chemical or physical properties of a compound based on existing knowledge of other compounds belonging to the same series.[137,138] Though less common, one can also infer structure from an observed chemical/physical property trend.

We recently demonstrated that mass – mobility correlations may be useful in identifying CH_2-homologous series in complex, carbon-rich samples such as bio-oils.[136] In particular, highly linear mass-to-charge (*m/z*) ratio vs. arrival time correlations were observed for several oxygen classes identified by Kendrick mass defect (KMD) analysis. We also observed dependencies of arrival time on oxygen class and double bond equivalents (DBE). Higher oxygen classes occupied smaller drift time windows compared to lower oxygen classes, and slopes of mass – mobility trendlines for compound classes with higher DBE were steeper than slopes of trendlines for compound classes with lower DBE.

At present, relatively few studies have focused on recognizing correlations between arrival time and chemical structure within homologous series. A unique linear correlation between carbon number and inverse reduced mobility ($1/K_o$) was reported by Hariharan *et al*. for several CH_2-homologous series representing primary and secondary

alcohols, ketones and aldehydes.[115,134] The same studies also demonstrated that such a correlation could be used to predict reduced mobilities of other missing members of a given series. Similar correlations between $1/K_o$ or drift time and *m/z* were observed by Demoranville et al. in the analysis of a homologous series of tetraalkylammonium ions.[139] Zamora *et al.* relied on correlation between carbon number and $1/K_o$ to tentatively identify an amidoamine series in coconut oil,[140] and Ahmed *et al*. investigated the correlation between carbon number and drift time to gain structural insight on sulfonated aromatic compounds in crude oil.[101] Recently, Katzenmeyer *et al.* interpreted a linear correlation between collision cross-section and *m/z* of fragments resulting from collision-induced dissociation of sodiated polylactide ions as an indication of linearity in the structure of the polymer.[141] Although each of these studies demonstrates the potential utility of using mass – mobility correlations in the analysis of homologous series, a systematic study focused on understanding the influence of chemical structure in such correlations has not been reported. Additionally, it is unclear whether linear correlations are universal among homologous series.

In the current study, several commercially-available homologous series were investigated with the aim of more systematically characterizing the influence of analyte structure on observed mass – mobility correlations. Experiments were designed to evaluate the effect of differing repeat units in polymeric structures. The identity of ionizable terminal groups and their position(s), relative to the propagating alkyl chain, were variables in the evaluation of several CH_2-homologous series. Inclusion of acidic and basic terminal groups afforded an evaluation of the potential effect of ion charge on mass – mobility trends. Series with aromatic and non-aromatic terminal groups were also

included. An interesting correlation between the slope of linear mass – mobility trendlines and DBE was observed for several CH_2-homologous series, providing new insight into the potential structure of alkyl chains present in select series, representing different oxygen classes, identified in a representative bio-oil.

Experimental Section

Chemicals and Reagents

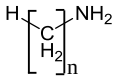
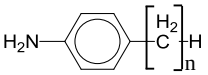
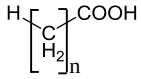
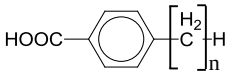
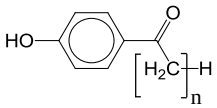
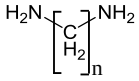
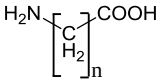
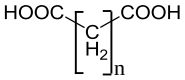
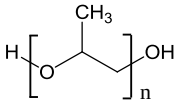
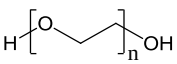
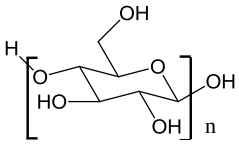
Unless noted otherwise, all chemicals were reagent grade or better, obtained from commercial vendors (Sigma-Aldrich and VWR Scientifics) and used as received. While PPG and PEG were purchased as a mixture, other compounds were purchased in "pure" form. HPLC grade methanol was purchased from Fisher Chemicals. Distilled water was purified and deionized to 18.2 MΩ with a Barnstead Nano-pure Diamond UV water purification system. All analytes, except glucose oligomers, were dissolved and diluted in methanol. Stock solutions of each compound (ca. 500 parts per million (ppm)) were prepared separately by dissolving 2 mg of compound in 4 mL methanol. Final solutions, containing all members of a given homologous series, were prepared by diluting an aliquot of each stock solution with methanol. Concentrations of different compounds in these mixtures were intentionally varied between 0.1 and 10 ppm so that moderate mass and mobility peak intensity of all components could be observed in a given spectra. Glucose oligomer solutions were prepared in a similar way, except water was used as the solvent. Slow-pyrolysis bio-oil was generated in a homebuilt reactor from corn-stover feedstock. Details pertaining to the preparation of this bio-oil are reported elsewhere [136]. Bio-oil was dissolved and diluted to a final concentration of 0.5 mg/mL in methanol

before analysis. All solutions were stored in air-tight glass vials at –20 ^{o}C when not in use. A complete list of compounds studied in the current analysis is provided in Table 3.1.

Data Acquisition and Analysis

All analyses were performed on a Synapt G2-S HDMS quadrupole–time-of-flight hybrid mass spectrometer (Waters Corp., Manchester, UK) coupled to a nano-electrospray ionization source (nano-ESI). IM-MS data were acquired with the time-of-flight analyzer operated in high-resolution mode ($m/\Delta m_{50\%}$ ~40,000 at *m/z* 300). The ESI source was operated in either negative or positive mode, depending on the propensity of a given compound to form negative or positive ions. Since it was feasible to ionize compounds in the amino-carboxlylic acid series as both deprotonated and protonated ions, this series was analyzed independently with the ESI source operated in both negative and positive modes. Instrumental parameters held constant during all analyses are given in Table 3.2. IM-MS data were processed using MassLynx software (Ver. 4.1). Raw mobility data were first smoothed using a Savitzky-Golay function with the "smooth window channel" and "number of smooths" each set equal to 2. The centroid of a given arrival time distribution was then determined as the arrival time corresponding to the midpoint of peak width when measured at half its maximum intensity. Best-fit lines presented in this work were obtained via linear regression of peak centroid data for all members of a given series. Data included in regression analyses represent six independent measurements of a given series made over a three-month time period.

Table 3.1. Commercially-available homologous series investigated in the current analyses

Series	Compound class	Structure	DBE	Number of repeat units (n)	Ion type
1	Amines		0	6–8, 10	$[M+ H]^+$
2	Anilines		4	2, 5–8, 10	$[M+ H]^+$
3	Carboxylic acids		1	4–6, 8	$[M – H]^–$
4	Benzoic acids		5	0–3, 5	$[M – H]^–$
5	Phenolic-ketones		5	1–8	$[M – H]^–$
6	Diamines		0	6, 7, 10, 12	$[M+ H]^+$
7	Amino-carboxylic acids (–)		1	4, 5, 7, 11	$[M – H]^–$
8	Amino-carboxylic acids (+)				$[M+ H]^+$
9	Dicarboxylic acids		2	2–4, 6	$[M – H]^–$
10	Polypropylene glycols			3–13	$[M+ Na]^+$
11	Polyethylene glycols			3–13	$[M+ Na]^+$
12	glucose oligomers			2–6	$[M+ Na]^+$

All statistical comparisons of slopes observed for different homologous series were performed using a two-tailed Student's *t*-test at the 95% confidence level. Kendrick mass defect analysis was carried out, as described previously,[34,136] to identify various CH_2-homologous series in the bio-oil sample. Double bond equivalents (DBEs) were calculated using a common equation.[28]

Table 3.2. Instrumental parameters held constant for all IM-MS analyses

Instrument Parameter	Value
ESI Infusion Rate (μL/min)	0.5
ESI Voltage (kV)	2.7
Source Temperature (°C)	90
Desolvation Temperature (°C)	150
Sampling Cone Voltage (V)	30
Extraction Cone Voltage (V)	35
Trap Voltage (V)	2.0
Transfer Voltage (V)	1.0
IM Wave velocity (m/s)	1000
IM Wave height (V)	40
IM Cell Pressure (mbar)	3.1
Helium Cell Pressure (mbar)	1.4×10^{3}
Trap Cell Pressure (mbar)	2.3×10^{-2}
Transfer Cell Pressure (mbar)	2.4×10^{-2}

Results and Discussions

Effect of Terminal Group Structure on Mass – Mobility Correlations Observed for CH_2-homologous Series

Representative raw data resulting from IM-MS analysis of a commercially-available CH_2-homologous series are shown in Figure 3.1. These data show the characteristic linear distribution of ions observed for a given series when analytical response is plotted with respect to mass and mobility dimensions. Note that different "non-homologous" species of similar *m/z*, such as the contaminants at nominal masses

150, 178, and 255, can be readily differentiated in the mobility dimension. The chemical structure in Figure 3.1 is representative of several "externally-propagating" series investigated in the present study. This terminology was chosen to emphasize that all compounds in these series have a single, terminal group (G) located at the end of a propagating alkyl chain (*i.e.*, H–$(CH_2)_n$–G). Homologous series comprised of compounds containing two terminal groups were also investigated (*i.e.*, G–$(CH_2)_n$–G′). The latter series are hereafter referred to as "internally-propagating" series. Graphs in Figure 3.2 were generated by plotting the centroid of extracted mobility data (Figure 3.1, left axis) versus nominal mass, number of repeat units, or number of atoms. These data clearly demonstrate that highly linear correlations ($R^2 \geq 0.996$) were observed for all CH_2-homologous series investigated in this work.

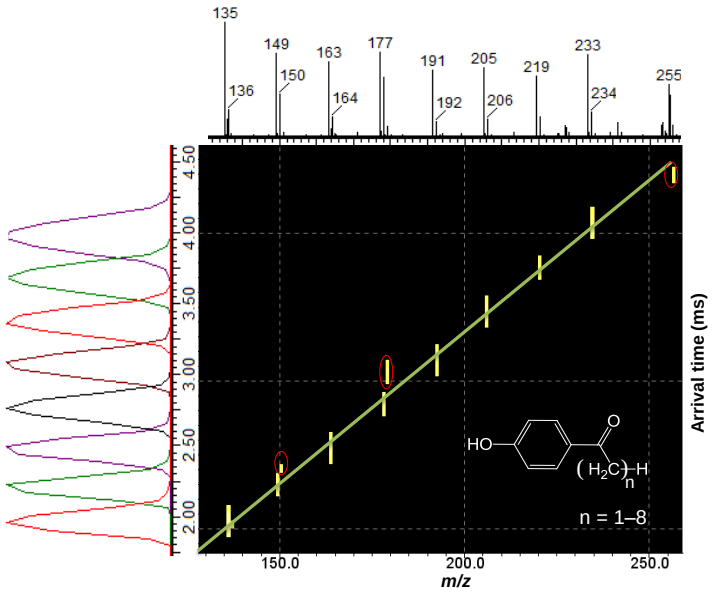


Figure 3.1. Representative ESI-IM-MS data for a CH_2-homologous series, showing the mass spectra and two-dimensional IM-MS spectra for the sample (top and central panels, respectively), as well as the extracted arrival time distribution (ATD) for each member of the homologous series (left panel). The trendline highlights the linear mass – mobility correlation observed for this series. Red ellipsoids in the IM-MS spectra designate peaks resulting from impurities.

Visual comparison of trendlines in Figure 3.2 led to the initial hypothesis that all series exhibiting a given structural motif (*i.e.*, all internally-propagating or externally-propagating series) would be expected to produce a similar slope. Regression equations, including uncertainties affiliated with the slope and intercept of each line, are reported in Table 3.3. A subsequent comparison revealed that the four slopes derived from internally-propagating homologous series (*i.e.*, Series 6-9) were statistically indistinguishable from one another (*i.e.*, the confidence intervals associated with the slope of each line overlapped). Three of five slopes corresponding to externally-propagating series were also statistically identical (*i.e.*, Series 1, 4, and 5). The slope observed for both Series 2 and Series 3 (*i.e.*, the remaining externally-propagating series) was statistically unique relative to the slope observed for any other series. However, all slopes corresponding to externally-propagating series were appreciably steeper than slopes observed for internally-propagating series.

The larger range of slopes observed for externally-propagating series, compared to internally-propagating series, may be related to variation in the relative size of terminal groups (G) represented in each structural motif. Among externally-propagating series that were ionized the same way (*i.e.*, as protonated or deprotonated ions), a larger slope was observed for series possessing a larger terminal group. That is, the slope observed for Series 4 and 5 was larger than that observed for Series 3, and the slope observed for Series 2 was larger than that observed for Series 1. Since the terminal groups represented in internally-propagating series investigated in this work are all expected to be of similar size, it is perhaps logical that these series produced very similar slopes in Figure 3.2. It follows that the range of slopes observed for internally-propagating series would be

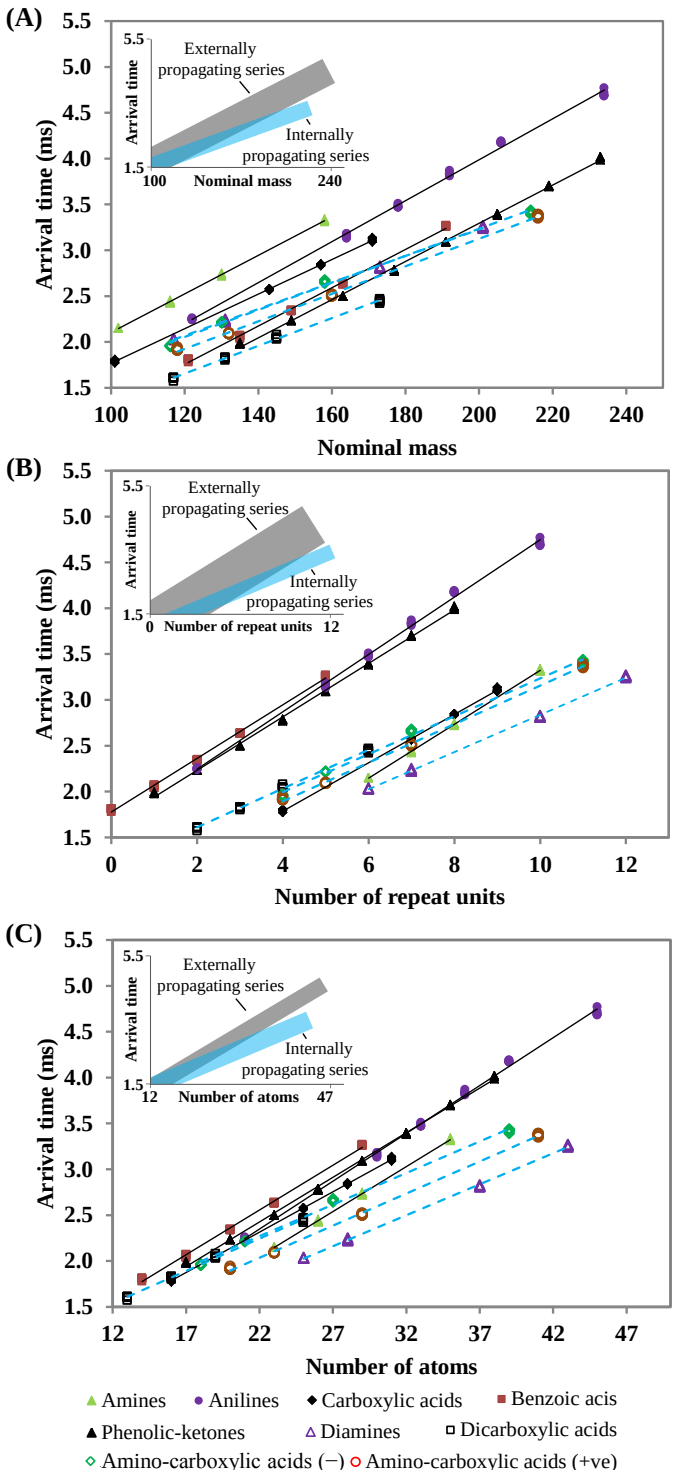


Figure 3.2. Arrival time plotted versus nominal mass (a), number of repeat units (b), and number atoms (c) for commercially-available CH2-homologous series. Space-filling plots (insets located in the upper left-hand corner of each graph) highlight the substantially different linear correlations observed for externally-propagating and internally-propagating series.

expected to increase if series with more variation in the relative size of terminal groups were included in the data set.

The inset located in the upper left-hand corner of each graph shown in Figure 3.2 provides an illustrative summary of how data reported in these plots trended in the context of the aforementioned hypothesis. Note that the range of coordinate space occupied by externally-propagating and internally-propagating series varied, depending on the identity of the variable plotted along the *x*-axis. It is apparent from comparison of Figure 3.2A-C that visual differentiation between externally-propagating and internally-propagating series in coordinate space is generally maximized at higher nominal mass, number of repeat units, and number of atoms. The convergence of coordinate space occupied by the two structural groups at lower values along the *x*-axis is consistent with increasingly limited conformational flexibility of smaller ions. However, data in Figure 3.2C suggest that plotting versus number of atoms may narrow the vertical range occupied by a given structural motif and minimize the degree of overlap between groups. Such minimization could be advantageous in the context of analyzing "unknown" mixtures. It is important to appreciate that high resolution mass spectrometry is required to accurately determine molecular formulas, defining number of atoms, in a complex sample.

Although the analysis presented in Figure 3.2 appears to be most useful in terms of distinguishing externally- and internally-propagating series, other factors were also considered in initial attempts to elucidate prevalent trends within collected data for commercial series. For example, data in Figure 3.2B suggest that series possessing a larger terminal group (*i.e.*, Series 2, 4, and 5 in this case) are likely to appear at longer

Table 3.3 Linear regression equations displaying uncertainty in the slope and intercept as a 95% confidence interval.[a]

Series	Compound class	Linear correlations			R^2
		Arrival time –nominal mass	Arrival time-number of atoms	Arrival time-number of repeat units	
CH_2-homologus series with externally propagating unit –CH_2– unit					
1	Amines	y = [0.0210 + 0.0002] x + [0.00 + 0.03]	y = [0.098 + 0.001] x – [0.11 + 0.03]	y = [0.294 + 0.002] x + [0.38 + 0.03]	1.000
2	Anilines	y = [0.0223 + 0.0004] x – [0.48 + 0.07]	y = [0.104 + 0.002] x + [0.06 + 0.06]	y = [0.313 + 0.005] x + [1.62 + 0.05]	0.998
3	Carboxylic acids	y = [0.0188 + 0.0002] x – [0.12 + 0.04]	y = [0.088 + 0.001] x + [0.38 + 0.03]	y = [0.264 + 0.003] x + [0.73 + 0.02]	0.999
4	Benzoic acids	y = [0.0209 + 0.0004] x – [0.75 + 0.06]	y = [0.097 + 0.002] x + [0.41 + 0.04]	y = [0.292 + 0.005] x + [1.78 + 0.01]	0.998
5	Phenolic-ketones	y = [0.0208 + 0.0002] x – [0.87 + 0.04]	y = [0.097 + 0.001] x + [0.29 + 0.03]	y = [0.292 + 0.003] x + [1.65 + 0.02]	0.999
CH_2-homologus series with internally propagating unit –CH_2– unit					
6	Diamines	y = [0.0145 + 0.0002] x + [0.33 + 0.03]	y = [0.068 + 0.001] x + [0.34 + 0.03]	y = [0.203 + 0.003] x + [0.81 + 0.03]	0.999
7	Amino-carboxylic acids (–ve)	y = [0.0147 + 0.0004] x + [0.28 + 0.06]	y = [0.069 + 0.002] x + [0.76 + 0.05]	y = [0.206 + 0.006] x + [1.17 + 0.04]	0.996
8	Amino-carboxylic acids (+ve)	y = [0.0150 + 0.0002] x + [0.12 + 0.04]	y = [0.070 + 0.001] x + [0.49 + 0.03]	y = [0.210 + 0.003] x + [1.05 + 0.02]	0.999
9	Dicarboxylic acids	y = [0.0151 + 0.0004] x – [0.16 + 0.05]	y = [0.071 + 0.002] x + [0.69 + 0.03]	y = [0.212 + 0.005] x + [1.19 + 0.01]	0.997
Oligomeric series[b]					
10a	Polypropylene glycols	y = [0.0123 + 0.0001] x + [0.29 + 0.04]	y = [0.079 + 0.001] x + [0.47 + 0.04]	y = [0.711x + 0.007] x + [0.79 + 0.04]	0.999
10b	Polypropylene glycols	y = [0.0172 + 0.0002] x − [2.2 + 0.1]	y = [0.111 + 0.001] x − [2.0 + 0.1]	y = [1.00 + 0.01] x − [1.5 + 0.1]	0.999
11a	Polyethylene glycols	y = [0.0104 + 0.0001] x + [0.52 + 0.01]	y = [0.065 + 0.001] x + (0.68 + 0.01]	y = [0.456 + 0.004] x + [0.94 + 0.02]	1.000
11b	Polyethylene glycols	y = [0.0137 + 0.0002] x − [0.65 + 0.08]	y = [0.086 + 0.001] x − [0.44 + 0.07]	y = [0.601 + 0.007] x − [0.09 + 0.07]	0.999
12	Glucose oligomers	y = [0.0119 + 0.0001]x + [0.21 + 0.06]	y = [0.0920 + 0.007]x + [0.42 + 0.02]	y = [1.93 + 0.02]x + [0.69 + 0.06]	0.998

a Regressions were performed using Microsoft Excel's (version 2010) LINEST function. Confidence intervals were calculated by multiplying the standard deviation of the slope or intercept (s_m or s_b, respectively) times the two-sided Student's t statistic for $N - 2$ degrees of freedom; $N = 6 \times$ the number of compounds in a given series. Note that the tabulated correlation coefficient (R^2) was identical for all three regression lines reported for a given homologous series.
[b] In each PPG and PEG series, letters "a" and "b" denote less steeper and more steeper trendlines, respectively.

arrival times for a given number of repeat units. The potential effect of charge state was also considered. When series possessed terminal groups of similar size, those forming protonated ions generally resulted in a more positive slope than series forming deprotonated ions (Table 3.3). For example, the slope observed for Series 2 was greater than that observed for Series 4 or 5. Similarly, the slope observed for Series 1 was greater than that observed for Series 3. Slopes calculated for series 7 and 8 also followed the same trend, but in this case the confidence intervals overlapped. Several studies have demonstrated a propensity for methanol to form adducts/clusters with positive ions.[142,143] Mass spectra and mobility data were carefully examined to determine whether cluster formation could be contributing to this observed trend in the present study. However, experimental evidence supporting such an argument was not found (see Appendix B, Figures B.1-B.3for details). An interesting correlation between slope and DBE was also observed.

As shown in Figure 3.3, a positive correlation between slope and DBE was generally observed for CH_2-homologous series analyzed in this study. This was first appreciated when slope data for commercially-available series were analyzed, but the strength of the correlation was leveraged significantly by data resulting from negative ESI-IM-MS analysis of a representative bio-oil. Bio-oil series plotted in Figure 3.3 were selected based on relative abundance (*i.e.*, several of these series are comprised of molecules representing the more abundant peaks observed in IM-MS spectra) and/or the feasibility of isolating, according to *m/z*, an arrival time distribution that qualitatively exhibited characteristics of a single gas-phase species. As documented in an earlier paper,[136] the general trend of increasing slope with increasing DBE was not unexpected

for bio-oil constituents. However, the correlation shown in Figure 3.3 also enables a novel interpretation of molecular structure in bio-oil. Notice that externally-propagating series are well separated from the trend formed collectively by internally-propagating and bio-oil series. This observation suggests that the CH_2-homologous series identified in bio-oil are structurally similar to internally-propagating commercial series. It is possible that the compounds present in bio-oil represent a linear alkyl chain sandwiched between two terminal groups. The bio-oil series may also be composed of structures containing a single terminal group and a propagating alkyl chain that is highly branched at the opposite end of the molecule. Additional experiments utilizing complementary techniques would be required to further detail the structure of molecules present in bio-oil. Nonetheless, data presented here provide an excellent example of how IM-MS may be useful in future efforts to more comprehensively characterize the chemical composition of bio-oils or other complex mixtures containing homologous series of compounds with a $(CH_2)_n$ repeat unit.

Mass – Mobility Correlations Observed for Oligomeric Series

Three oligomeric series were also analyzed by IM-MS. As demonstrated in Figure 3.4, a linear correlation was observed for a series composed of glucose repeat units. In contrast, a break in linearity was observed for oligomeric series composed of ethylene glycol and propylene glycol repeat units. That is, an initial linear relationship between arrival time and mass was observed up to a specific number of repeat units (7 for PEG and 8 for PPG). Beyond this point, an independent linear correlation exhibiting a steeper slope was observed as additional repeat units were added. This observation was rationalized by considering potential gas-phase coordination chemistry involving sodium

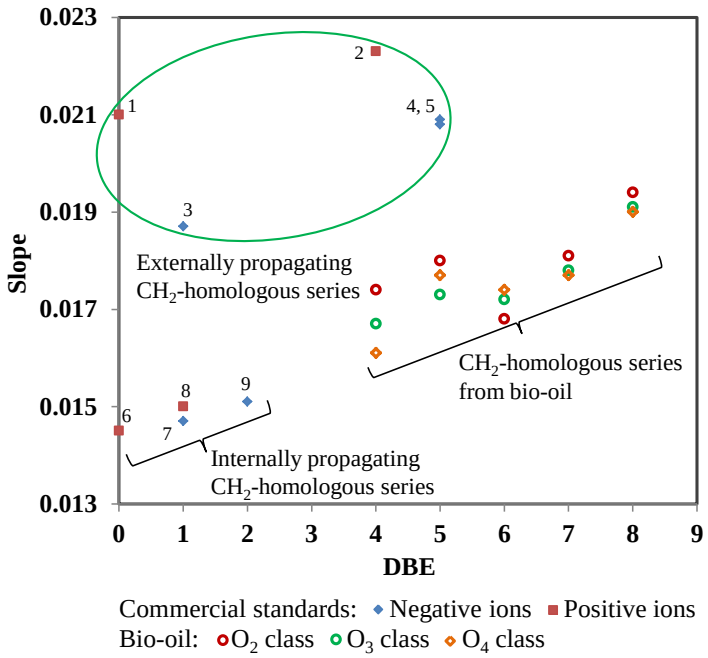


Figure 3.3. Scatter plot of slopes resulting from mass – mobility correlations plotted vs DBE for various commercial series and select CH_2-homologous series identified in bio-oil. Note the general increase in slope with increasing DBE. Additionally, the contiguous trend formed by internally-propagating series and the series identified in bio-oil suggests structural similarity. Numbers alongside the markers for commercial series indicate the series number as it appears in the Table 3.1.

ions. Previous molecular dynamics calculations performed by Gidden et al. on sodiated forms of PEG and PPG suggest that the sodium ion interacts with up to 8 oxygen atoms.[144] Thus for PEG oligomers, it is possible that oxygen present in each repeat unit from n = 3 to n = 7 coordinates with a central sodium ion forming a relatively compact structure. If additional repeat units do not interact with sodium, they would be expected to contribute more conformational freedom to the ion and produce a larger change in collision cross section per added repeat unit beyond n = 7. A similar argument can be made for PPG oligomers.

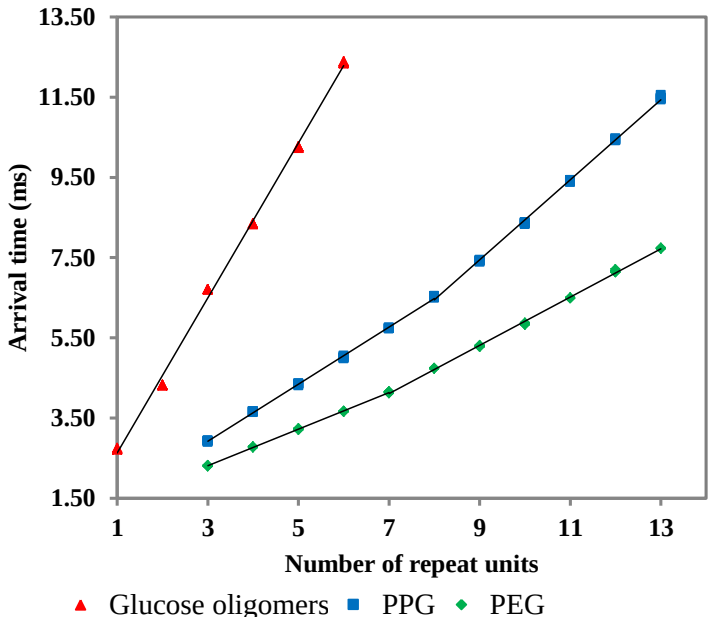


Figure 3.4. Arrival time versus number of repeat units plots for oligomeric series composed of polyethylene glycols (PEG) polypropylene glycols (PPG) and glucose oligomers. A break in linearity at n = 7 in PEG and n = 8 in PPG may be indicative of a substantial change in the gas-phase structure of sodiated ions.

The analysis of oligomeric series was initially performed in order to evaluate potential effects of varying the structure of repeat units in homologous series. For the series investigated in this work, a clear trend involving slopes observed for various repeat units was difficult to discern in correlations of arrival time with nominal mass (*m/z*) or number of atoms (see Table 3.3). In contrast, slopes trended in the following order when arrival time was correlated with number of repeat units: CH_2 < ethylene glycol < propylene glycol << glucose. This trend is likely a consequence of the relative number of atoms present in each repeat unit (CH_2 = 3, ethylene glycol = 7, propylene glycol = 9, glucose = 21) and/or the relative mass of each repeat unit (CH_2 = 14, ethylene glycol = 44, propylene glycol = 58, glucose = 162), both of which may be roughly correlated with size.

Conclusions

Systematic variations of slope observed in this work suggest that linear mass – mobility correlations can not only support grouping of compound classes, but also tentative identifications of a specific structural motif in IM-MS spectra of a complex sample. Such a statement is supported by the correlation shown in Figure 3.3, suggesting that select CH_2-homologous series in a representative bio-oil are structurally similar to internally-propagating commercial series. Additionally, results observed for PEG and PPG oligomers demonstrate that changes in mass – mobility correlations within a given homologous series may also provide structural insight into relevant gas-phase coordination chemistry. One could easily envision developing a "library" of slope values corresponding to a wide range of structural variation that could be used to develop algorithms that facilitate rapid structural profiling in complex mixtures. It is important to point out that arrival time data reported in the present study are likely to vary for alternative instrumentation and/or experimental parameters (*i.e.*, gas pressure, wave height, wave velocity, etc.). Thus, future work in this area should seek to include determinations of collision cross section (CCS) for investigated homologous series. Strategies for measuring CCS *via* traveling-wave IMS have been previously reported,[145,146] and the availability of CCS data should enable deduction of slope values that are less dependent on experimental conditions. It will also be important to evaluate alternative structural motifs (*e.g.* alternative terminal groups or the effect of alkyl branching), as well as homologous series encompassing a larger mass range.

CHAPTER FOUR

Characterization of Bio-oil Solubility Fractions Using High-resolution Mass Spectrometry and Ion Mobility Spectrometry

Abstract

The oily phase of slow-pyrolysis bio-oil generated from pine shavings biomass was subjected to sequential solvent fractionation for separating constituent compounds by their polarities. Solvents used for fractionation expriments include hexane, water, diethyl ether and dichlormomethane. Bio-oil and resulting solvent fractions were subsequently analyzed using high-resolution mass spectrometry and ion mobility – mass spectrometry (IM-MS) in negative-ion mode electrospray ionization. Analyte ion signals were observed primarily between *m/z* 80 and 500 in the mass spectra. Hundreds of oxygen-rich compounds containing 1 to 11 oxygen atoms per molecule were identified. Partial separation and enrichment of compounds by oxygen classes was observed in various solvent fractions where oxygen classes generally trended with solvent polarities. The utility of IM-MS was evaluated to for the potential separation for isomeric compounds present in the bio-oil. IM-MS analysis of bio-oil demonstrated solvent fractionation-like separation of oxygen classes. Additionally, potentially isomeric compounds present in these solvent fractions were evaluated by comparing arrival time and collision-induced dissociation data of select ions.

Introduction

Pyrolysis of biomass produces numerous decomposition products of lignocellulosic compounds in the form of pyrolysis vapors, aerosols and solid residue.[11] The liquid obtained by the condensation of pyrolysis vapors and aerosols is called bio-oil, and it is an attractive feedstock for use as fuel or value-added chemicals. Bio-oil is a complex mixture comprising varieties of oxygenated and highly polar compounds that are produced as a result of many simultaneous and sequential reactions.[13–15] While there is increasing interest in utilizing bio-oil as a feedstock for producing transportation fuels, development of suitable upgrade techniques have been the major challenge.[7] However, before upgrade strategies can be developed, better understanding of molecular composition of bio-oils is crucial.[121]

Conventional techniques of bio-oil analyses are limited to either bulk properties (*e.g.*, functional group identification of the bulk bio-oil by nuclear magnetic resonance[17,21,22] or Fourier transform infrared spectroscopy[18–20]) or identification of volatile compounds (*e.g.*, by gas chromatography-mass spectrometry; GC-MS[12,22–24]). Recently, high- and ultrahigh-resolution mass spectrometry (HRMS) has been sought for more comprehensive analysis of bio-oils.[31,48–50,127] Although the majority of these analyses were performed using Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) which is capable of providing the highest resolving power and mass accuracy,[147] very high purchase and operating costs limits the use of this technique for routine analysis. Results from studies involving mass spectrometers of different resolving power suggested that Orbitrap with with resolving power ~30,000 at *m/z* 400 can be an reasonable choice for routine analysis of bio-oils.[48,148] HRMS analyses

of bio-oils have enabled the identification of hundreds to thousands chemical formulas of bio-oils' constituents and have greatly improved understanding of bio-oil composition. However, it is important to note that the structural elucidation of bio-oil constituents using only MS is challenging without a complementary separation, specifically in a compositionally complex sample that consist of potentially isomeric compounds.[31,32]

Studies involving sequential fractionation of bio-oil with solvents of varying polarities followed by analyses using gel permeation chromatography (GPC) and GC-MS have been reported.[15,22,24] These studies demonstrated that compounds of bio-oils can be separated into several "macro-chemical families". However, detailed understanding of composition of resulting solvent fractions has not been accomplished. Liu *et al.* utilized FT-ICR MS to analyze such solvent fractions of a fast-pyrolysis generated bio-oil from red-pine feedstock and demonstrated enrichment of different compound classes in different solvent fractions.[50] However, as noted by these authors, compounds with the same molecular formulas were present with substantially different relative abundance in spectra of these solvent fractions. While it is possible that same molecular formulas (*i.e.*, *same m/z* peaks) observed in different solvent fractions are the result of partitioning of the compounds during fractionation, it is also possible that they represent isomeric compounds. Therefore, further analysis using techniques capable of separating isomers would be expected to provide futher details. Additionally, compositions of bio-oils generated using different feedstocks and/or different pyrolysis conditions can be substantially different. For example, relatively low molecular weight distributions[10] and less compositional complexity[136] of slow pyrolysis bio-oil compared to the fast pyrolysis bio-oils have been reported. Therefore, analysis of bio-oils generated under different

thermochemical conditions (*e.g.*, slow versus fast pyrolysis) and/or from different feedstock types could be useful in improving the overall understanding of bio-oils composition.

In this study, slow-pyrolysis bio-oil generated from pine shavings was subjected to fractionation, using hexane, water, ether, and dicholomethane (DCM), for separating bio-oil constituents according to their polarities. Compositions of bio-oil and resulting solvent fractions were subsequently analyzed using negative-ion mode ESI-Obrbitrap MS. In addition, isomeric compounds potentially present in these solvent fractions were investigated by comparing arrival time and collision-induced dissociation (CID) data of select ions. This work represent the first application IM-MS in combination with HRMS for the compositional analysis of various solvent fractions of bio-oil derived using traditional fractionation technique.

Experimental Section

Preparation of Slow-pyrolysis Bio-oil

Details of the preparation method of slow-pyrolysis bio-oil has been described in previous publication from our group.[136] Briefly, ~1.5 kg of pine-shavings feedstock was heated in a custom-built airtight bio-char reactor (a 20 L stainless steel reaction vessel). Biomass was heated to a maximum temperature of 490 °C at an average heating rate of ~4.5 °C/min where the total reaction time, including the initial temperature ramp, was 200 min and oil condensation was accomplished at ambient air temperature (~9 °C). Two bio-oil phases, oily and aqueous, were obtained by phase separation. In the current work, only the oily phase was subjected to solvent fractionation and subsequent analyses.

Solvent Fractionation

Sequential solvent fractionation was achieved using a similar procedure as previously described,[22,50] and the scheme of the fractionation is illustrated in Figure 4.1. To 50 grams of bio-oil, 200 gram of n-hexane was added in a round bottom flask. The mixture was first stirred for 2 hours using a magnetic stir bar and then shaken for 2 hours using MaxQ 4000 shakers (Thermo Scientific, San Jose, CA) at the speed of 500 rpm. The hexane soluble fraction was separated using a separatory funnel and the solvent was removed using a vacuum rotary evaporator at 40 °C. Similar steps were used for further fractionation with other solvents (*i.e.*, water, diethyl ether and dichloromethane (DCM) except DCM soluble and insoluble, as well as ether soluble and insoluble fractions were separated using filters of pore size 0.1 μm (Pall Corporation), rather than using a separatory funnel. A sample-to-solvent ratio was maintained at ~1:40 in each step. Solvents used in this analysis were of HPLC grades or higher quality (purchased from Sigma-Aldrich).

Data Acquisition and Analysis

The sample solutions (~1 mg/mL) for negative-ion mode ESI-Orbitrap and negative-ion mode ESI-IM-MS analyses were prepared by dissolving 5 mg of each sample separately in 5mL methanol. Full-scan mass spectra in the *m/z* range 50 to 600 were obtained with ESI operated in negative-ion mode on a linear ion trap-Orbitrap mass spectrometer (LTQ-Orbitrap Discovery, Thermo Scientific, San Jose, CA). A sample solution was injected into the ESI source at 10 μL/min using a syringe pump (Hamilton 500 μL) and data were acquired with Orbitrap mass spectrometer operated at a mass resolving power (R) of 30,000 (at *m/z* 400; R = m/ Δ m50%). Optimized source

conditions were as follows: sheath gas flow, 60; auxiliary gas flow, 5; sweep gas flow, 5; transfer capillary voltage, 8 V; tube lens, 80 V; capillary temperature 250 °C; electrospray needle voltage 4 kV.

For each sample, data were collected for 200 scans with a maximum injection time of 300 ms. Prior to analysis, each sample was spiked with sodium dodecyl sulfate (SDS) (final concentration 0.1 ppm) and the lock mass was set on the accurate mass of the deprotonated SDS ion ($[M - H]^-$) at *m/z* 265.1479 for internal mass calibration. With this strategy, high mass accuracy (*e.g.*, average RMS error < 1.3 ppm) was generally observed. Accurate masses of spectral peaks having a relative intensity greater than 0.01% were exported from Thermo Xcalibur (Ver. 2.2) to a spreadsheet for subsequent data analyses.

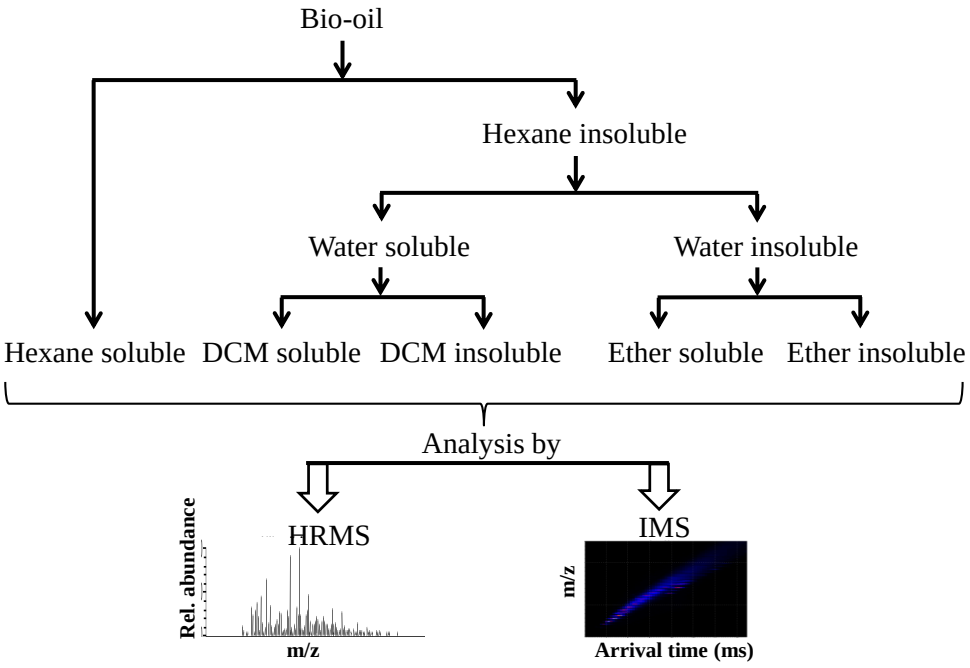


Figure 4.1. Bio-oil solvent fractionation and analysis scheme.

IM-MS and all CID data were obtained on a Synapt G2-S HDMS instrument (Waters Corp., Manchester, U.K.) equipped with a nano-ESI source operated in negative-mode. Parameters used during IM-MS analyses are given in Table 4.1. IM-MS/MS experiments involved selection of nominal mass of interest followed by IM separation and subsequent fragmentation in the transfer cell. During IM-MS/MS experiments, the transfer cell and wave velocity were set to 40 V and 2000 m/s, respectively. IM-MS data were acquired with the TOF analyzer operated in high-resolution mode (R = m/ Δ m50% ~40,000 at *m/z* 300). IM-MS data were processed using massLynx software (Ver. 4.1).

Table 4.1. Instrumental parameters used in IM-MS analyses

Instrument Parameter	Value
ESI Infusion Rate (μL/min)	0.5
ESI Voltage (kV)	2.7
Source Temperature (°C)	90
Desolvation Temperature (°C)	150
Sampling Cone Voltage (V)	30
Extraction Cone Voltage (V)	35
Trap Voltage (V)	2.0
Transfer Voltage (V)	1.0
IM Wave velocity (m/s)	1000
IM Wave height (V)	40
IM Cell Pressure (mbar)	3.1
Helium Cell Pressure (mbar)	1.4×10^{3}
Trap Cell Pressure (mbar)	2.3×10^{-2}
Transfer Cell Pressure (mbar)	2.4×10^{-2}

Results and Discussions

Molecular Weight Distribution and Compositional Analysis of Bio-Oil and Its Solvent Fractions

Representative (−)ESI-high-resolution mass spectra of bio-oil and its five solvent fractions analyzed in the current study are shown in Figure 4.2. Analyte ion signal was observed primarily between *m/z* 80 and 500 in the mass spectra of bio-oil and its solvent fractions. We had observed a similar molecular weight distribution of bio-oil generated from the same feedstock but at slightly different pyrolysis conditions (*e.g.*, biomass heating rate at 2.5 °C/min in previous study compared to 4.5 °C/min in the current study). Although similar molecular weight distributions were observed in the spectra of bio-oil and various solvent fractions, relative abundances of compound types varied substantially. For example, the ions with highest relative intensity are different in the spectra of various solvent fractions. Based on previously reported GC-MS results, potential structures of these most abundant ions are shown in Figure 4.2.

Weighted average molecular weight of bio-oil, hexane soluble, ether soluble, ether insoluble, DCM soluble and DCM insoluble fractions are, respectively, 202, 211, 217, 248, 178 and 181. Similarly, weighted median molecular weight of bio-oil, hexane soluble, ether soluble, ether insoluble, DCM soluble and DCM insoluble fractions are, respectively, 193, 203, 207, 241, 163 and 161 Da. Similar molecular weight distributions were observed when these samples were analyzed using negative-ion mode ESI-TOF-MS (Figure C.1).The observed central distribution of molecular weights in various solvent fractions in the current analysis, however, are much lower than previously reported mean molecular weights (~400 Da) of a fast pyrolysis bio-oil from red pine by Liu *et al.*, using

ultrahigh-resolution MS analysis.[50] Although we do not know the reason(s) for these observed differences, we speculate potential contribution of one or more of the factor(s) including; potentially different biomass storage and treatment conditions, different pyrolysis conditions (slow vs fast pyrolysis), different instruments and/or instrumental conditions (Orbitrap vs FT-ICR). However, it is important to note that the observed lower molecular weight distributions in this HRMS study compared to results from HRMS of fast pyrolysis by Liu *et al* are consistent with previous observations (using traditional techniques) and proposed thermochemical reaction mechanisms of pyrolysis.[10]

A representative expanded region at a single nominal *m/z* 295 in Figure 4.3 demonstrates a visible shift of base peak toward higher mass defect or higher oxygenated compound from hexane fraction to DCM fraction. Since polarities of these solvent are in the order hexane < ether < DCM, the increased relative intensities of higher oxygenated compounds in the spectra of higher polarity solvents fraction may suggest separation (or enrichment) of bio-oil constituents by their polarities. Additionally, the presumed increased concentration of specific compound classes and resulting higher relative abundance mass spectra may justify qualitative treatment of observed relative intensities of peaks in negative-mode ESI-Orbitrap mass spectra.

It is important to note that some ions which are not observed in the bio-oil itself are present with a good S/N ratio in solvent fractions. For example, the peak assigned with chemical formula $C_{13}H_{11}O_8$ is observed in ether soluble fraction and DCM insoluble fraction but not in bio-oil. This is potentially due to reduced matrix effect on ESI[149] of these compounds and/or enrichment of these compounds in these solvent fractions.

Similar trends were also observed at many other mass regions. In certain cases, it was also observed that a peak appearing only as a shoulder of a nearby larger peak, or hidden completely under it, appears as a prominent peak in one or more solvent fractions. A

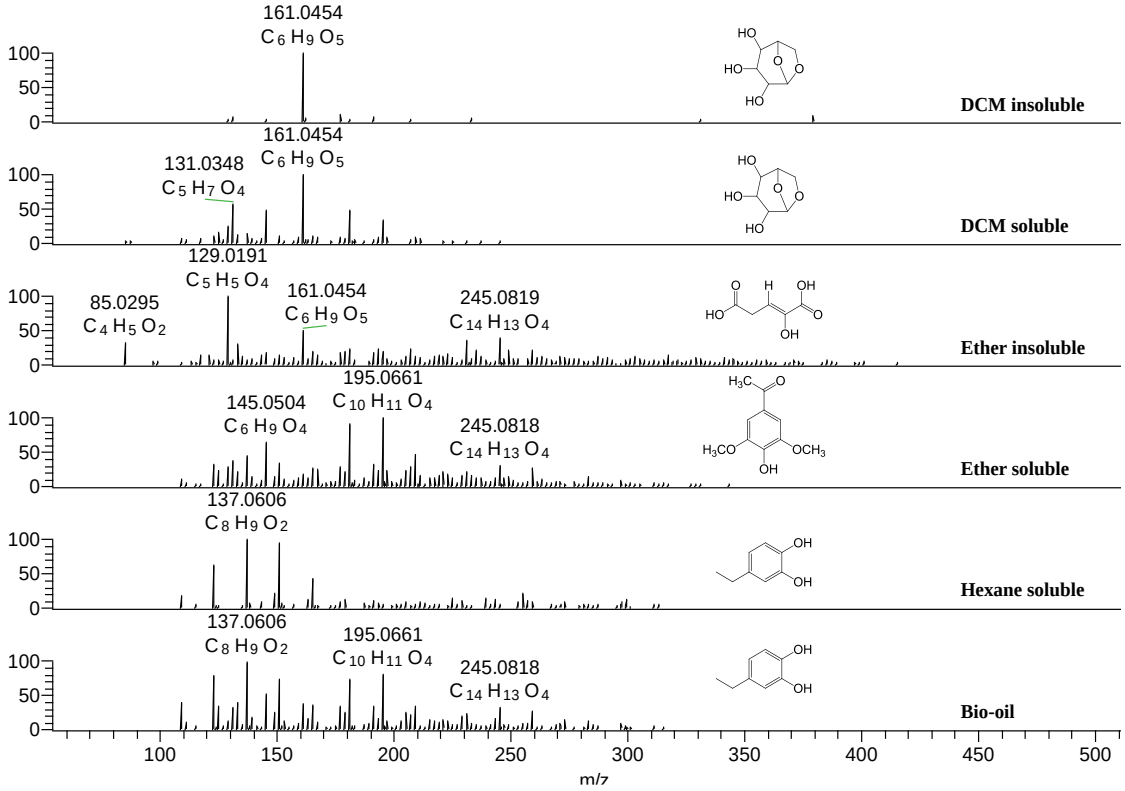


Figure 4.2. Molecular weight distributions observed in (−)ESI-LTQ-Orbitrap mass spectra of bio-oil and its various solvent fractions. Potential chemical structures of the most abundant ion in each spectrum is illustrated in the figure.

representative example is illustrated in the Figure 4.4. It is also observed that, in such cases, seemingly odd or no chemical formula was suggested by data processing software, Xcalibur. For example, $C_5H_{15}O_{11}N_2$ was the only formula suggested for a peak with *m/z* = 279.0680 in DCM soluble fraction under study-imposed criteria of elements (*i.e.*, C $\leq$ 40, H $\leq$ 80, O $\leq$ 30 and N $\leq$ 2). These observations suggest that fractionation

could be beneficial for the comprehensive analysis, specifically when analyzed using Orbitrap or TOF mass spectrometer with resolving power ~30,000 at *m/z* 400.

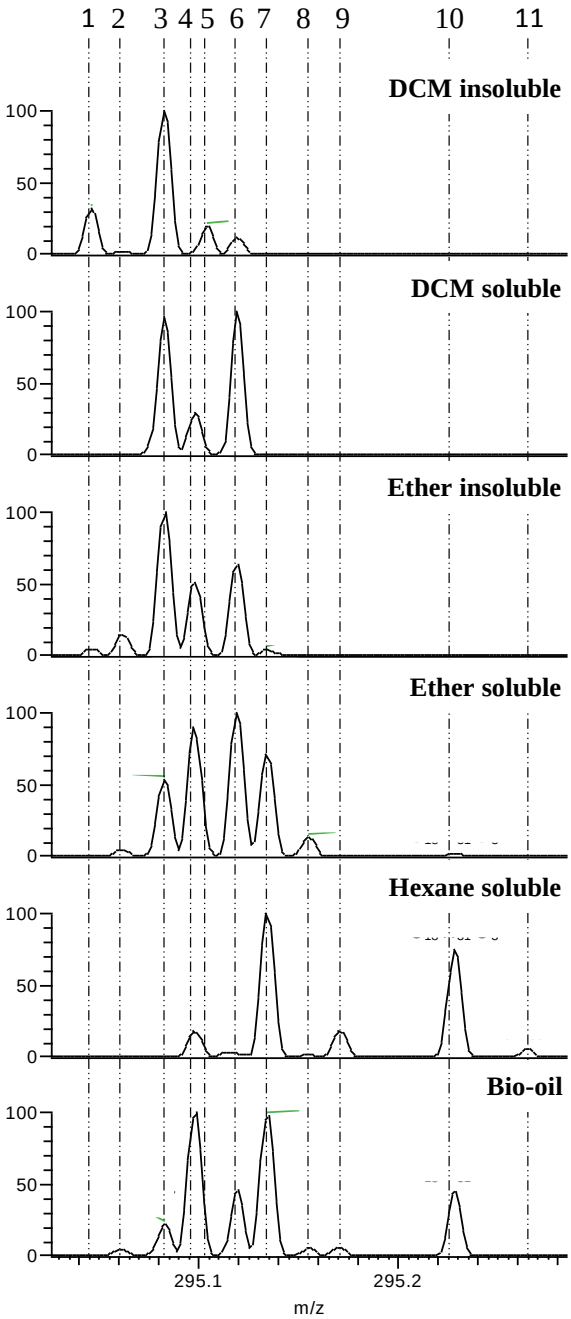


Figure 4.3. Expanded region of mass spectra at *m/z* 295. Selective enrichment of different compound classes in different solvent fractions is observed.

Table 4.2 *m/z* values and chemical formulas of mass spectra peaks shown in Figure 4.3

Peak	*m/z*	Formula
1	295.0458	$C_{13}H_{11}O_8$
2	295.0611	$C_{17}H_{11}O_5$
3	295.0825	$C_{14}H_{15}O_7$
4	295.0976	$C_{18}H_{15}O_4$
5	295.1037	$C_{11}H_{19}O_9$
6	295.1188	$C_{15}H_{19}O_6$
7	295.1340	$C_{19}H_{19}O_3$
8	295.1550	$C_{16}H_{23}O_5$
9	295.1704	$C_{20}H_{23}O_2$
10	295.2279	$C_{18}H_{31}O_3$
11	295.2642	$C_{19}H_{35}O_2$

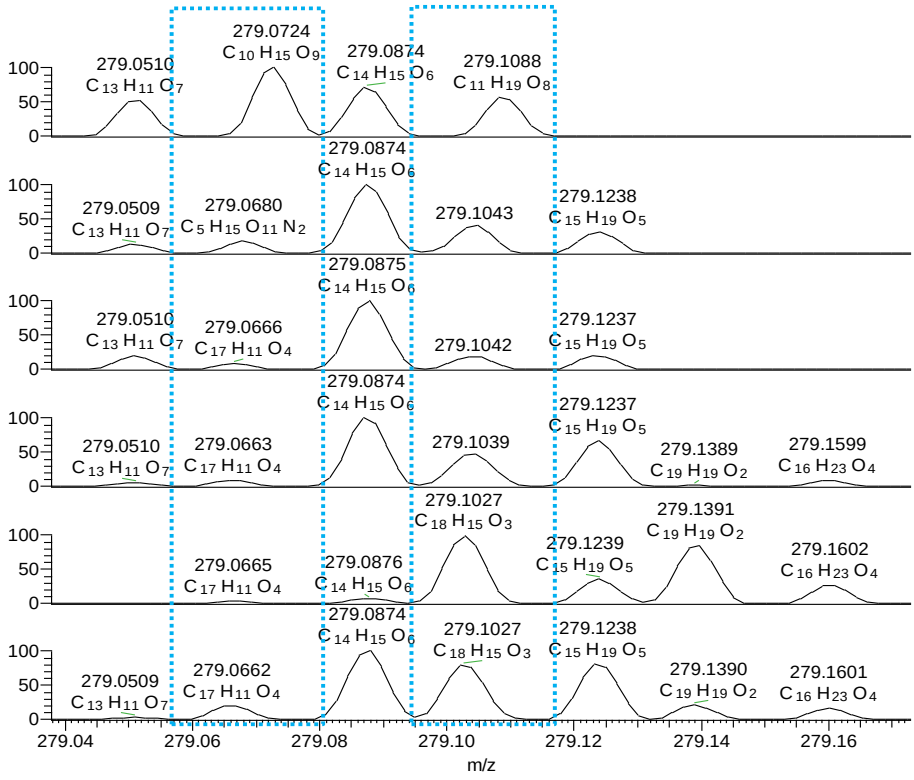


Figure 4.4. Expanded region of mass spectra of bio-oil and its solvent fractions at *m/z* 279 illustrating significance of fractionation for comprehensive analysis of bio-oil.

van Krevelen Diagrams

van Krevelen diagrams are useful for the tentative classification of compound classes based on the oxygen-to-carbon (O/C) and hydrogen-to-carbon (H/C) elemental ratios. Plots shown in Figure 4.5 are generated by plotting such ratios as observed in the molecular formulas of all identified compounds in the current work. The overlaid ellipsoids in the figure (in orange color) indicate spaces different compound classes would occupy in a van Krevelen plot. Majority of the identified compounds in bio-oil and its various solvent fractions exhibit O/C ratios between 0.1 and 1.0 and H/C ratios between 0.6 and 2.0. These plots demonstrate that compounds resembling condensed hydrocarbons are abundant in all but DCM soluble and DCM insoluble fractions. Whereas few carbohydrate-like compounds are much more abundant in DCM soluble and DCM insoluble fractions, lignin-like compounds appear to concentrate more in ether soluble and ether insoluble fractions. The selective extraction of compounds with lower O/C ratio in hexane fraction may suggest that bio-oil compounds in this solvent fraction could be superior to other solvent fractions from fuel prospective.

Relative Abundance and Distribution of Oxygen Classes

Relative abundance of various oxygen classes observed in bio-oil and its solvent fractions are shown in Figure 4.6. While O_4 class is the most abundant in bio-oil, and three of the solvent fractions, namely ether soluble, ether insoluble and DCM soluble, O_2, and O_5 compound classes are the most abundant in hexane soluble and DCM insoluble fractions, respectively. Nearly symmetrical relative abundances of various oxygen classes are observed in ether soluble and insoluble fractions, as well as DCM soluble and insoluble fractions. However, distribution of oxygen classes in bio-oil and hexane soluble

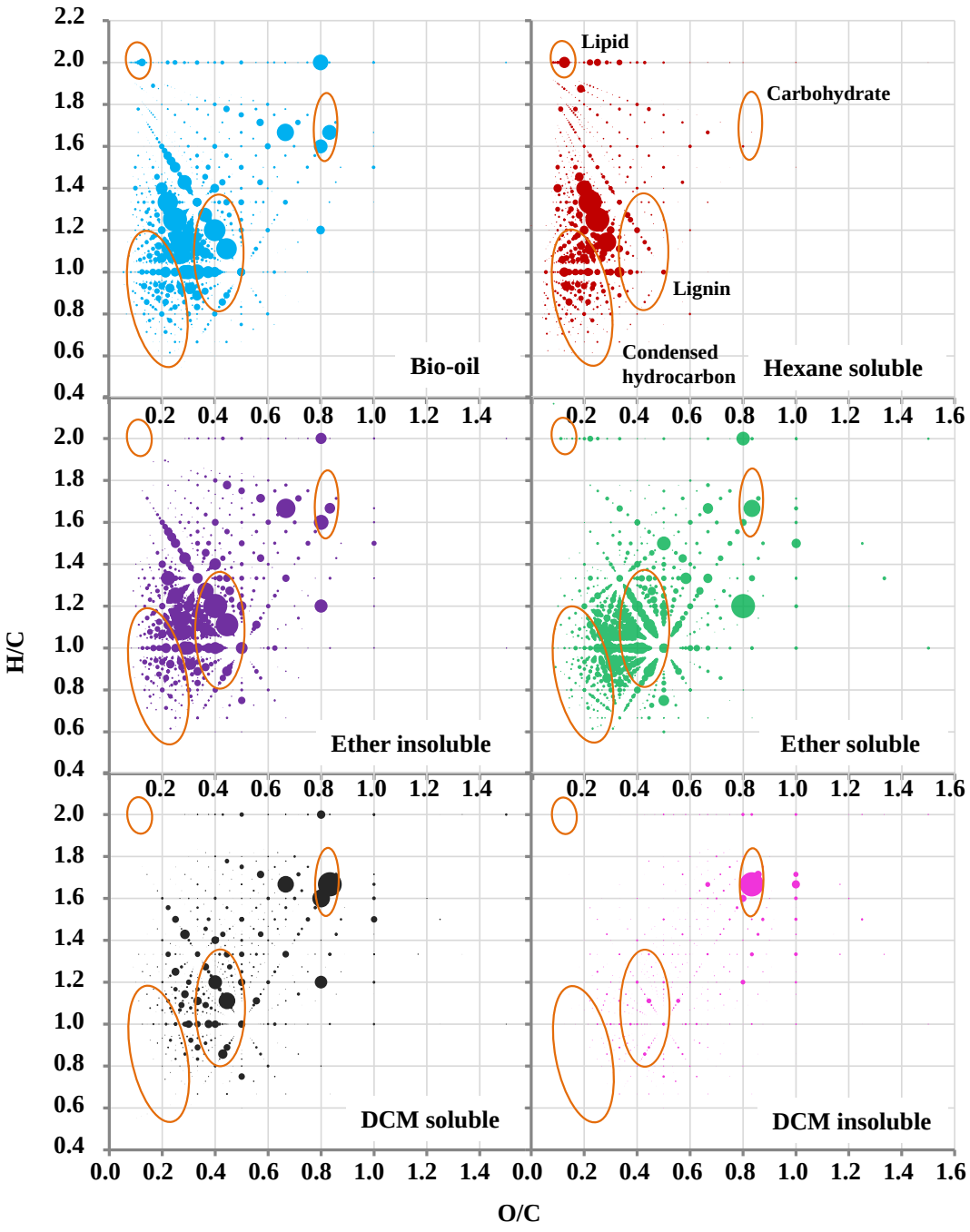


Figure 4.5. van Krevelen plot of bio-oil and its solvent fractions. Bubble size represents the relative intensity of corresponding ion in the spectra. To avoid crowding in the figures, regions representing different compound classes in van Krevelen plots are labelled only in the plot of hexane soluble fraction.

fractions are skewed towards lower oxygen classes. Some solvent fractions (*e.g.*, DCM insoluble fractions) appear to comprise higher oxygen classes not observed in the bio-oil sample. The highest oxygen class observed in the bio-oil is O_8 but up to O_{11} compound classes were observed in DCM insoluble fraction. Although reduced matrix effect on analyte ionization could result in similar observation of signal enhancement, observed results (*i.e.*, increased relative intensity of potentially more polar higher oxygen classes in the mass spectra of more polar solvent fractions) are more consistent with selective enrichment of compounds by their polarities during fractionation. It is possible that those higher oxygen classes are present in bio-oil at very low concentration levels and the resulting signal intensities are below study-imposed threshold (*i.e.*, relative intensity < 0.01) in the mass spectra. In that case, these compound classes would not be included in the subsequent data analyses. However, due to enrichment of these compounds in specific solvent during fractionation, their resulting spectral intensities are higher than the threshold and thus appear in the subsequent data analyses.

Contour plots of oxygen classes O_2 to O_8 of bio-oil and various solvent fractions observed in the current analysis are shown in Figure 4.7. Enrichment of specific compound classes in a given solvent fraction is more visible in these plots. As discussed in previous sections, oxygen classes generally trended with the solvent polarities, and these figures provide molecular-level comparison of compound types. Generally one or more CH_2-homologous series have the highest relative abundance in each oxygen class. A clear exception to this can be observed in O_5 class in both DCM soluble and insoluble fractions where, rather than a compound series, a single compound of formula $C_5H_{10}O_5$ is relatively much more abundant than others. Single model distribution with one broad

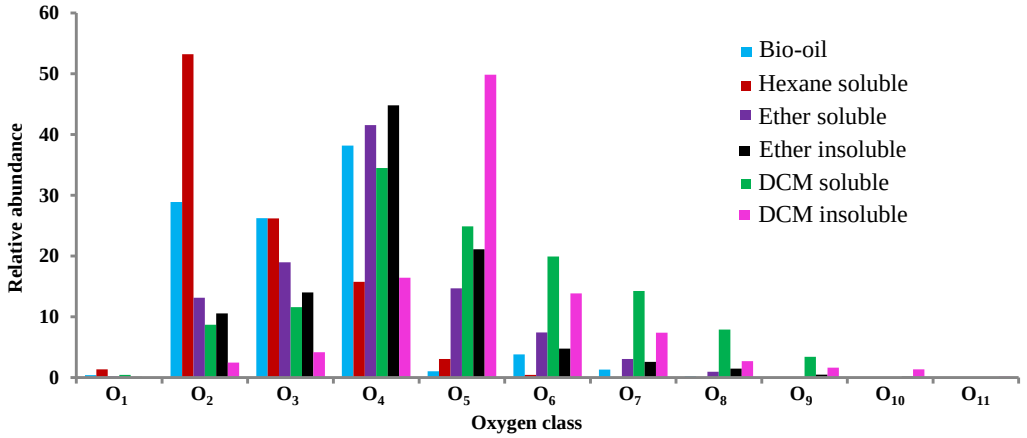


Figure 4.6. Relative abundance of various oxygen classes observed in bio-oil and its solvent fractions.

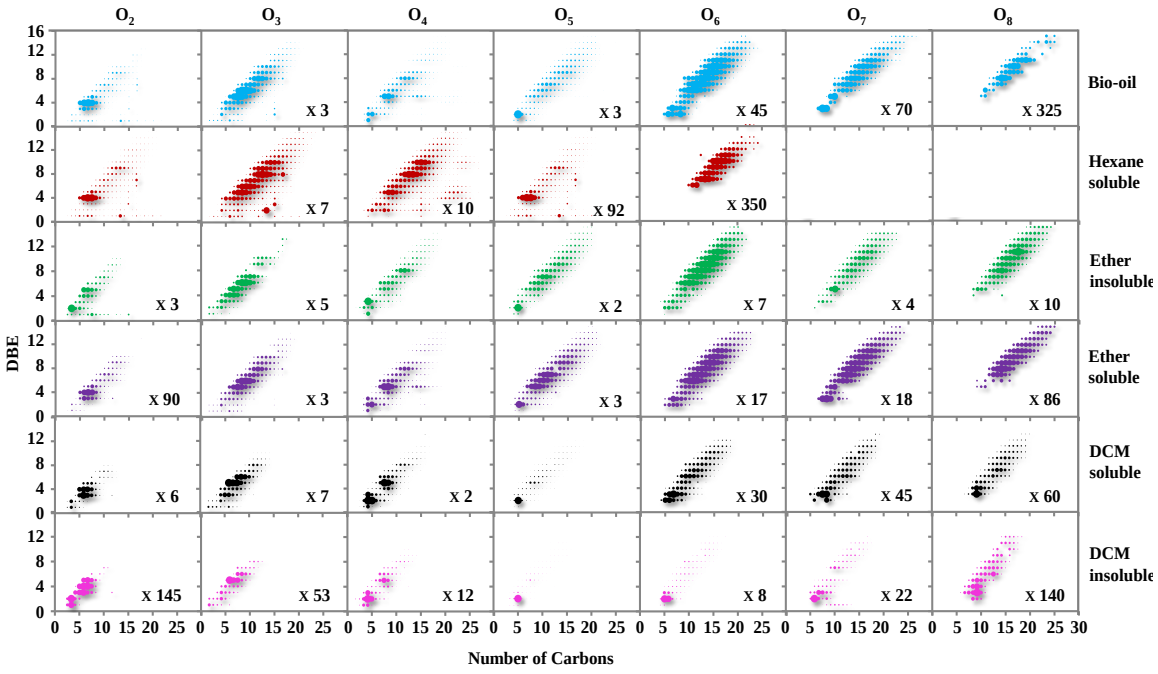


Figure 4.7. Contour plots of various oxygen classes observed in the bio-oil and its solvent fractions. Bubble size represents the relative intensity of corresponding ion in the spectra and the numbers represent the magnification scale applied to data points in the respective figures.

roughly linear trend was observed in majority of these plots with a few exceptions, *e.g.*, O_2 to O_4 in bio-oil and O_2 to O_5 in hexane solvent fractions. Observed results of single model distributions are consistent with relatively less selective reaction mechanism involved in the thermochemical decomposition of lignocellulosic micromolecules.

It is important to note a generally similar molecular weight distribution with a substantial difference in ion intensities in a particular oxygen class among different solvent fractions. Although the solvent-to-sample weight ratio used for the fractionation in the current study was higher than in similar previous studies,[22,50] the presence of many common *m/z* peaks in two or more solvent fractions may suggest an incomplete fractionation. Another explanation for the observed results could be the presence of isomeric compounds of substantially different polarities.

IM-MS Analysis

Representative two-dimensional (2D) ion mobility spectra of bio-oil and its solvent fractions are shown in Figure 4.8. It is observed that the majority of ions fall along a broad, roughly linear distribution in each spectrum. As noted in our previous study,[136] such a distribution is indicative of presence of similar compound classes in the sample. However, a careful comparison of ion distributions and resulting broad trendlines reveal somewhat gradual shifts of position toward lower space from hexane to ether to DCM fractions. For example, distribution of ions in hexane soluble and DCM insoluble fractions are visibly different from each other with majority of ions occupying the higher space in hexane fraction and lower space in DCM insoluble fraction. As discussed in previous sections (see Figure 4.6 and 4.7), above 90% of compounds identified in hexane fraction were O_4 or lower oxygen classes while above 90% of compounds in DCM

insoluble fractions are O_4 or higher oxygen classes. This observation is consistent with previously reported results by our group, where we demonstrated three CH_2-homologous series belonging to O_2, O_3 and O_4 classes occupied different 2D space.[136]

To further understand the potential separation of compound classes by IM-MS, we regenerated mass spectra of bio-oil from three regions of 2D plot as indicated in Figure 4.9A. Regenerated mass spectra (Figure 4 9B) and representative expanded regions at nominal *m/z* 269 (Figure 4 9C) demonstrate that the compounds occupying these regions belong to different classes. While the majority of compounds in the Region 1 are of lower oxygen classes (*e.g.*, O_4 or lower), compounds occupying Region 2 are primarily of higher oxygen classes (*e.g.*, O_5 and higher). If we compare compound classes observed in these regions with compounds identified in various solvent fractions, Regions 1 and 2 roughly resemble hexane soluble and DCM insoluble fractions, respectively, suggesting solvent fractionation-like separation of of bio-oil constituents by IM-MS.

We also processed data to obtain arrival time of hundreds of ions. CID experiments were performed on few post mobility-separated ions. Careful comparison of arrival time data and CID spectra of common *m/z* peaks present in two or more solvent fractions did not produce clear evidence of presence of isomeric compounds in different solvent fractions. However, it is important to note that not all isomeric compounds are expected to have different arrival time and/or CID spectra. Therefore, based on the results from current work, possibility of presence of isomeric compounds in different solvent fractions may not be ruled out.

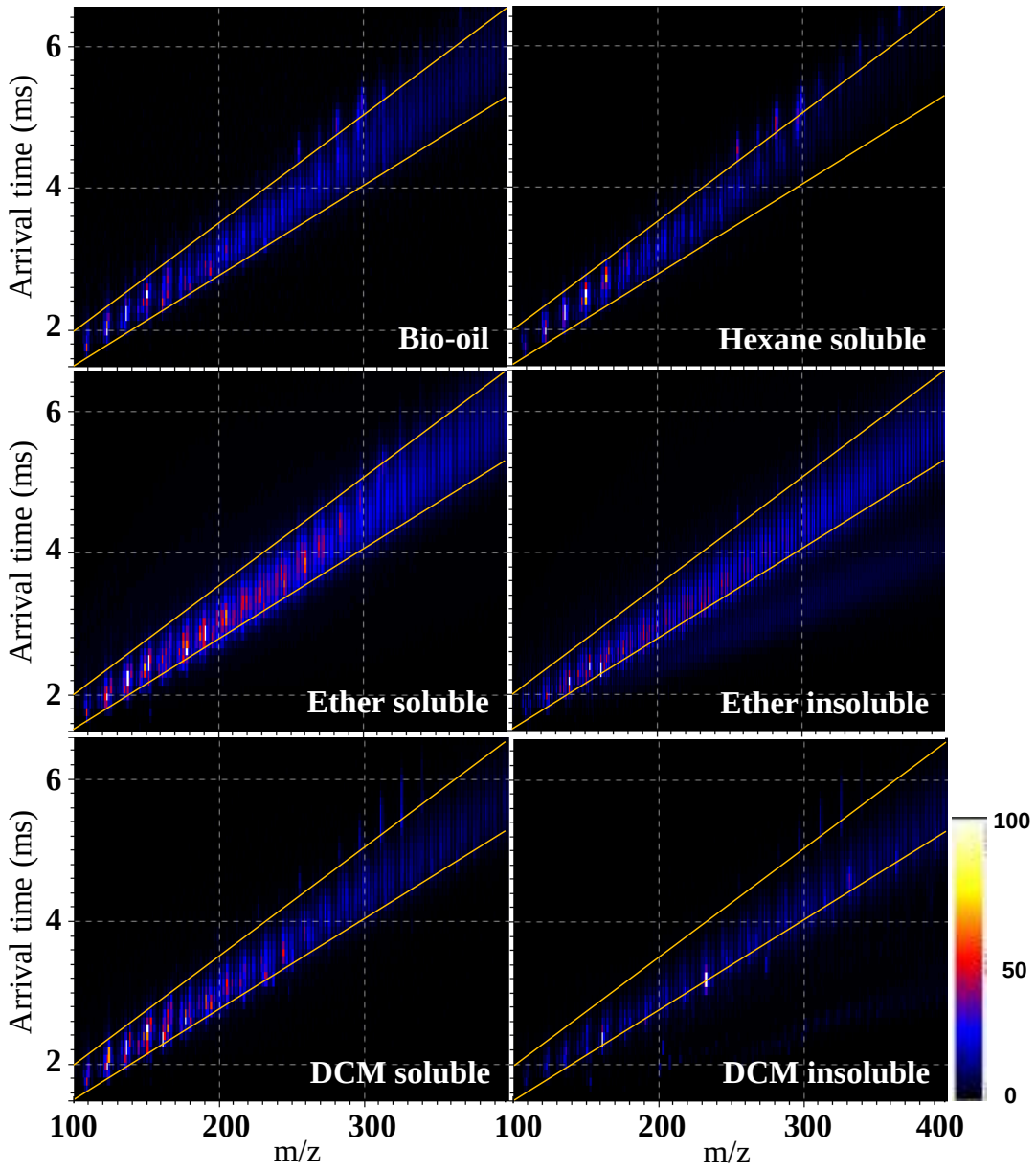


Figure 4.8. Two dimensional IM-MS spectra showing molecular distribution (between *m/z* 100 and 400) in bio-oil and its solvent fractions. Relative abundance of peaks are represented by coloration. Each spectrum is overlaid with two lines (in orange color) in exact same position to facilitate the comparison of distribution of ions in 2D space among different sample types.

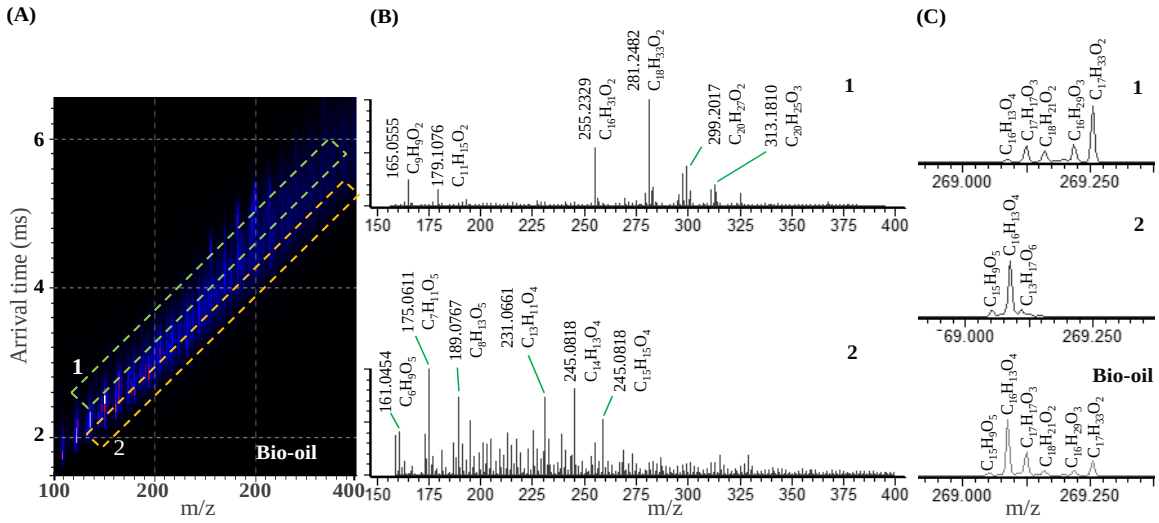


Figure 4.9. Rectangular Regions 1 and 2 (represented by green and orange rectangles, respectively) in 2D plot of bio-oil (Panel A) and resulting mass spectra from these select regions (Panel B) demonstrate that it is possible to separate different oxygen classes in bio-oil by IM-MS. Expanded regions at nominal mass 269 (Panel C) provide a closer look at compound classes observed in Regions 1 and 2.

Conclusions

HRMS and IM-MS analyses of PS bio-oil subjected to sequential solvent fractionation using solvent of increasing polarities suggest that partial separation and enrichment of compounds by oxygen classes can be achieved. Results from current study also demonstrated that solvent fractionation reduces the sample complexity and provide more comprehensive understanding of bio-oil composition. The partial separation of various oxygen classes of bio-oil in two dimensional space of IM-MS spectra roughly resembled solvent fractionation. This may suggest that with a higher resolving power in mobility dimension, IM-MS analyses might offer similar level of understanding of bio-oil composition as provided by sequential fractionation and HRMS analyses combined. It is important to note that solvent fractionation is a much rigorous task that requires hours to days to complete while IM-MS analysis can be performed in much shorter time scale (*e.g.*, few minutes).

CHAPTER FIVE

Conclusions and Future Directions

HRMS Analysis of Slow pyrolysis Bio-oil

Results from HRMS analyses of PS and CS bio-oils generated using slow pyrolysis (in Chapter Two) suggest that these liquids are generally similar in composition to previously reported fast-pyrolysis bio-oils. The second order mass defect analysis (MD^2) suggests a high degree of compositional similarity in terms of the more abundant components ($S/N \geq 15$) identified in aqueous and oily fractions of CS and PS bio-oils. Results from HRMS analysis of bio-oil and various solvent fractions described in Chapter Four suggest that bio-oil constituents are separated in different solvents according to oxygen class. In general, more polar solvents result in enrichment of higher oxygen classes. Accordingly, traditional solvent fractionation techniques are likely to be beneficial for comprehensive analysis of bio-oils; especially when the analysis relies on high-resolution (*e.g.*, resolving power ~30,000 at *m/z* 400), as opposed to ultrahigh-resolution mass spectrometry >200,000) .

It is important to note that all ionization techniques have a bias and result in preferential ionization of certain compound classes over others. For example, negative-ion mode ESI primarily produces deprotonated ions from acidic molecules and, in contrast, positive-mode ESI produces protonated ions from basic molecules.[82] Relatively less polar compounds such as hydrocarbons are less amenable to ESI and would not be observed when analyzed using this ionization technique. Therefore, the total composition

of complex samples such as bio-oils, which potentially contain several classes of compounds, may appear different when analyzed using different ionization techniques. For example, several previous analyses using GC-MS have identified saturated and unsaturated hydrocarbons, as well as compounds belonging to the O_1 oxygen class. Such compounds are either not detected (*e.g.*, hydrocarbons) or are present at very low relative intensity (*e.g.*, O_1 oxygen class) in negative-ion mode ESI mass spectra.[25,31] Therefore, it is important to analyze complex samples using complementary ionization techniques for more comprehensive sample characterization. Although a few HRMS studies have demonstrated the potential utility of other ionization techniques (*e.g.*, APCI and APCI) for analysis of bio-oils, these studies represent a "proof of concept", only highlighting potential applications of these other ionization techniques.[52,53] However, studies focused on improving the overall understanding of bio-oil composition resulting from different biomass types and pyrolysis conditions have not been reported. Future work in this field would seek to utilize different ionization techniques to gain a better understanding of bio-oil composition.

In the near future, we plan to analyze bio-oils using complementary ionization techniques, such as APCI and APPI for analysis of relatively less polar hydrocarbons which may be present in bio-oils. Preliminary data suggest a substantially different bio-oil composition (*e.g.*, with respect to observed constituents and/or relative intensity) relative to what was observed using ESI. Figure 5.1 compares expanded regions of ESI and APCI (both in negative-ion mode) mass spectra of bio-oil and its hexane-soluble fraction. Increased relative intensities of O_1 and O_2 compound classes and decreased

relative intensity of the O_4 compound class in APCI spectra compared to ESI spectra can be observed. These results suggest that APCI and ESI preferentially ionize compounds

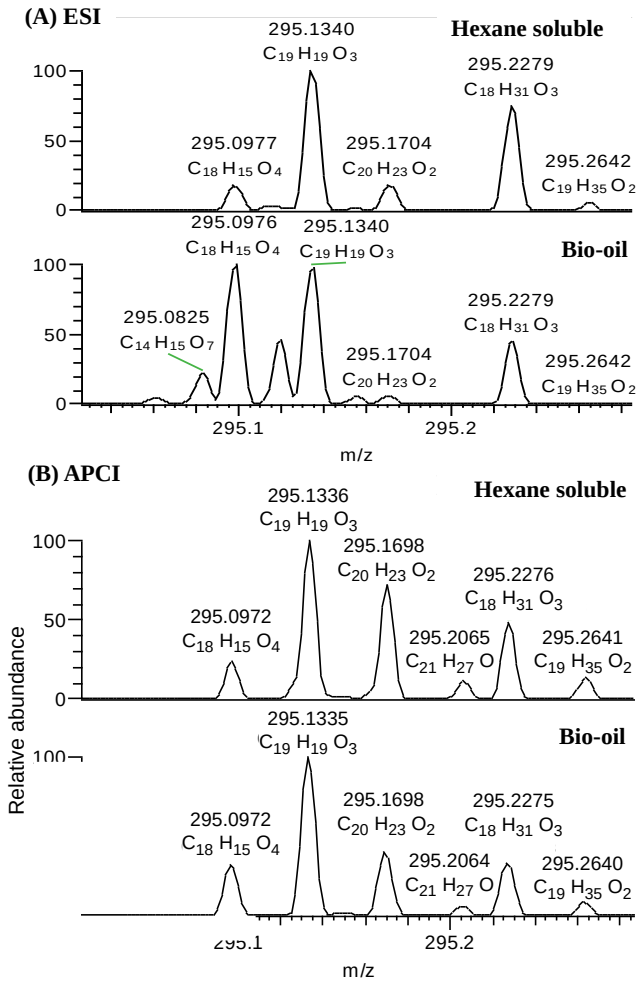


Figure 5.1. Expanded region at *m/z* 295 of (A) (−)ESI and (B) (−)) APCI -LTQ-Orbitrap mass spectra of bio-oil and its hexane soluble fraction. ESI data shown in this figure are reproduced from Chapter Four (Figure 4.3) for comparison to APCI spectra.

belonging to lower and higher oxygen classes, respectively. Similar observations are also made in van Krevelen plots (Figure 5.2) and relative abundance plots (Figure 5.3). van Krevelen plots demonstrate much higher density of data points in low O/C ratio regions in APCI spectra compared to ESI spectra suggesting detection of a larger number of compounds with a low O/C ratio (*i.e.*, low oxygen class). Moreover, relative abundance

plots show that O_1 compound classes are much more abundant in APCI spectra compared to ESI spectra.

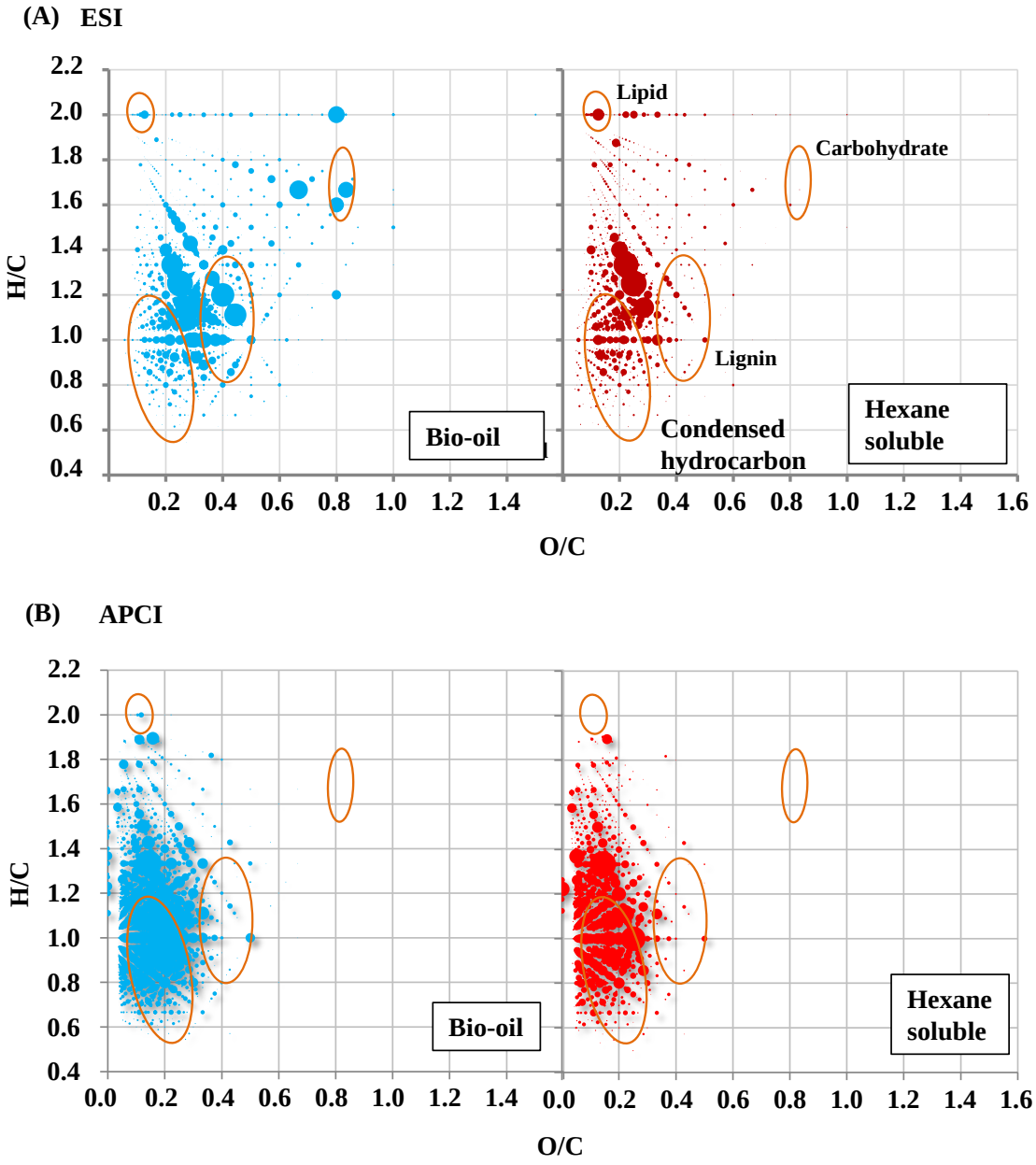


Figure 5.2. van Krevelen plots of bio-oil and hexane fraction plotted from (A)(−)ESI and (B) (−) APCI spectra. Bubble size represents the relative intensity of corresponding ion in the spectra. ESI data shown in this figure are reproduced from Chapter Four (Figure 4.5) for comparison.

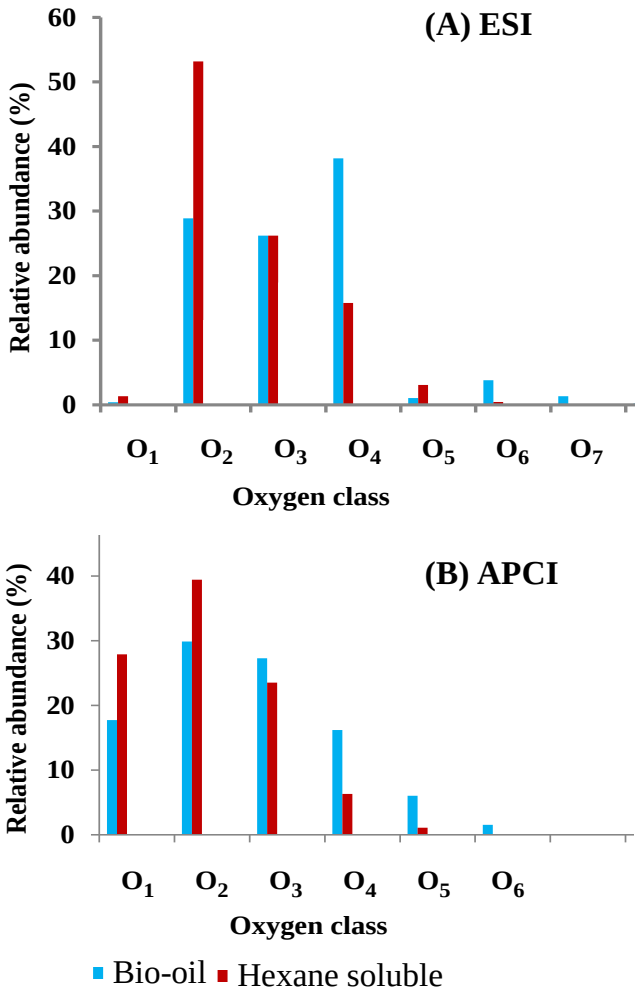


Figure 5.3. Relative abundance of various oxygen classes in bio-oil and hexane fraction from (A) (−)ESI and (B) (−)APCI spectra. ESI data shown in this figure are reproduced from Chapter Four (Figure 4.6) for comparison.

Future work should also focus on improving selectivity of fractionation to enhance the characterization and identification of various compound classes present in bio-oils. We consider two potential approaches for optimizing selectivity: (1) optimizing conditions and solvents during extraction and (2) selective pulling out compounds comprising specific functionality using highly-selective membranes, adsorbents, or reactivity trends.

IM-MS Analysis

Results from the current work (described in Chapter Two and Chapter Three) suggest that linear mass – mobility correlations can be useful for grouping similar compound classes, as well as identifying a specific structural motif in a complex sample. For example, our results indicated that select CH_2-homologous series in a representative bio-oil were structurally similar to internally-propagating commercial series. Additionally, results observed for PEG and PPG oligomers demonstrated that changes in mass – mobility correlations within a given homologous series could also provide structural insight into relevant gas-phase coordination chemistry. Studies focused on identifying mass – mobility correlations eventually may lead to developing a "library" of slope values corresponding to a wide range of structural variation that could be used to develop algorithms that facilitate rapid structural profiling in complex mixtures. It is important to point out that arrival time data reported in the present study are likely to vary for alternative instrumentation and/or experimental parameters (*i.e.*, gas pressure, wave height, wave velocity, etc.). Thus, future work in this area should seek to include determinations of collision cross section (CCS) for investigated homologous series. Strategies for measuring CCS *via* traveling-wave IMS have been previously reported,[145,146] and the availability of CCS data should enable deduction of slope values that are less dependent on experimental conditions. It will also be important to evaluate alternative structural motifs (*e.g.* alternative terminal groups or the effect of alkyl branching), as well as homologous series encompassing a larger mass range.

APPENDICES

APPENDIX A

Characterization of Slow-pyrolysis Bio-oils by High-resolution Mass Spectrometry and Ion Mobility Spectrometry

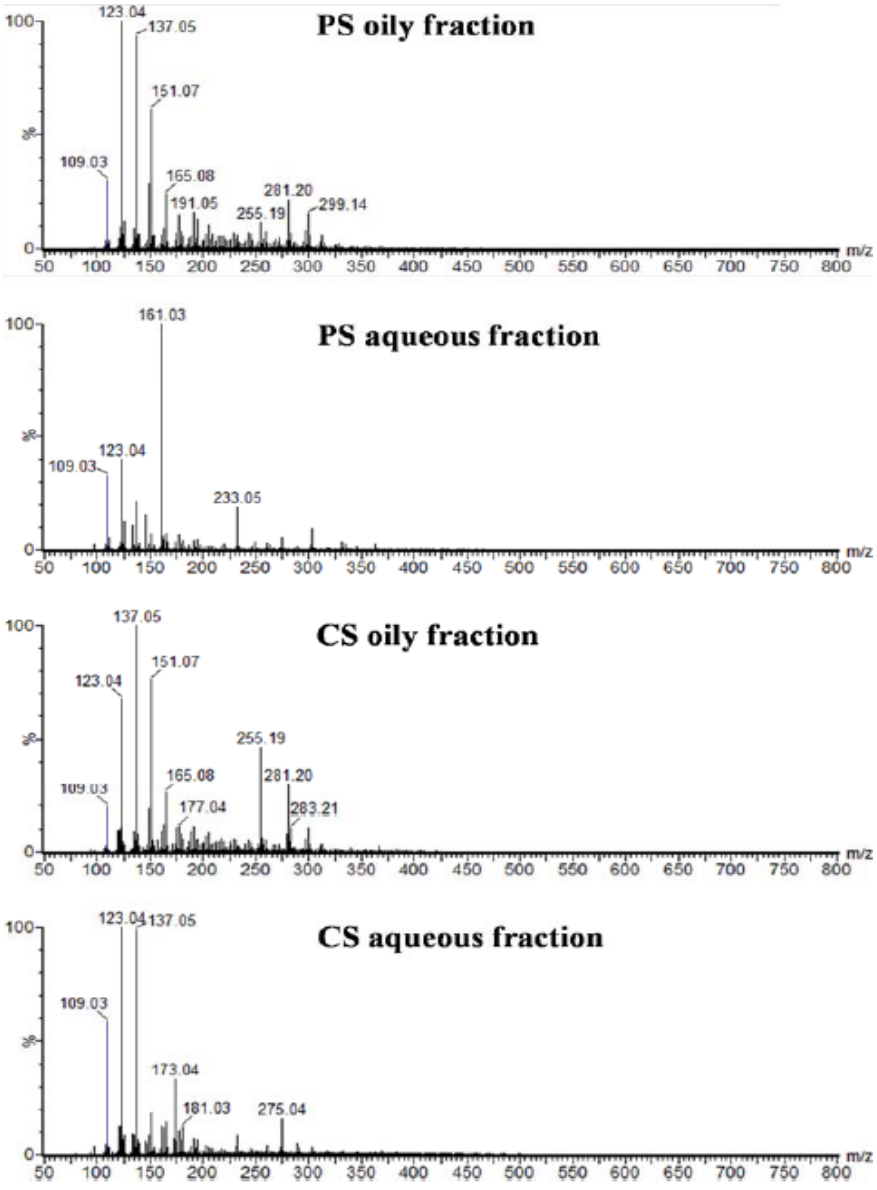
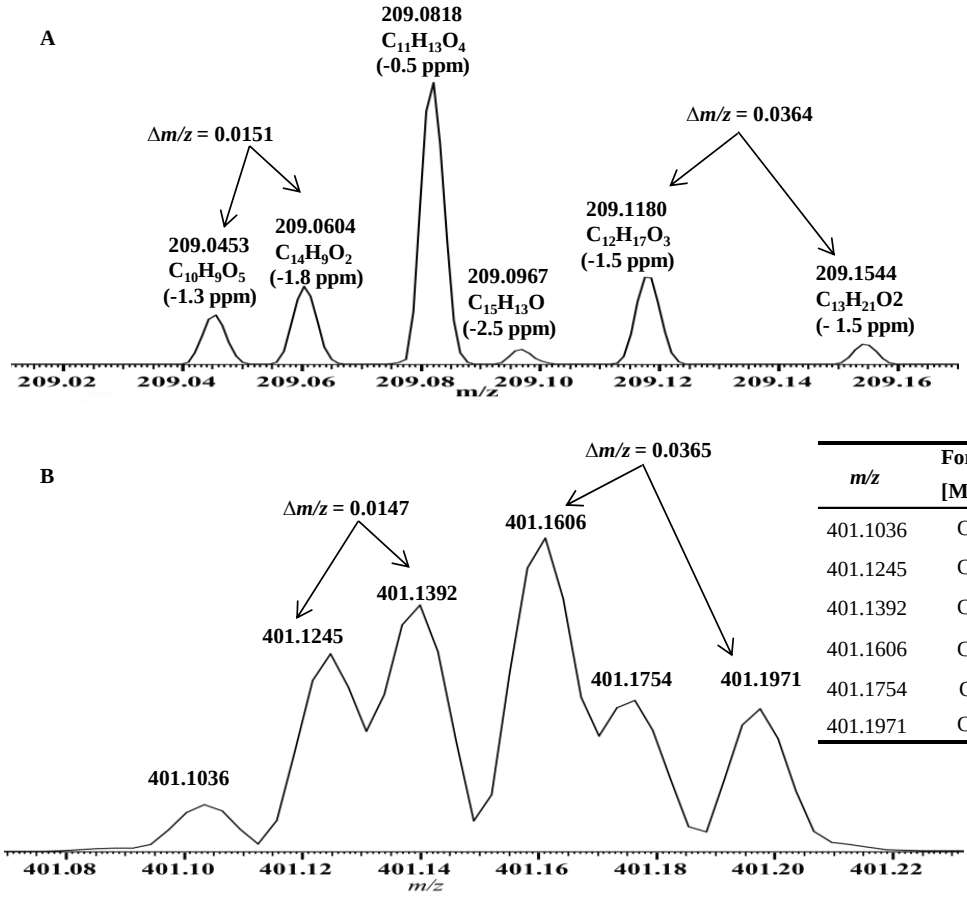
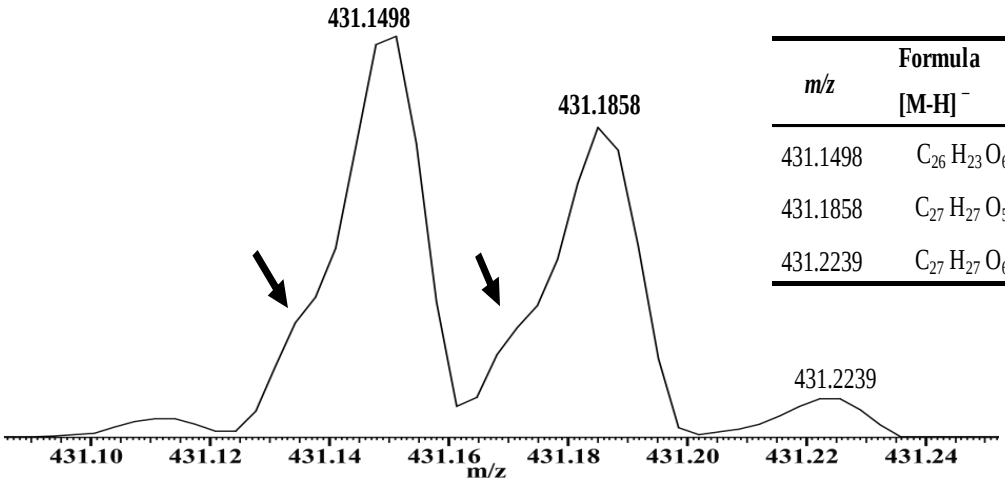


Figure A.1. Molecular weight distributions of the four bio-oil fractions as observed by (–)ESI-TOF-MS (Synapt G2-S).



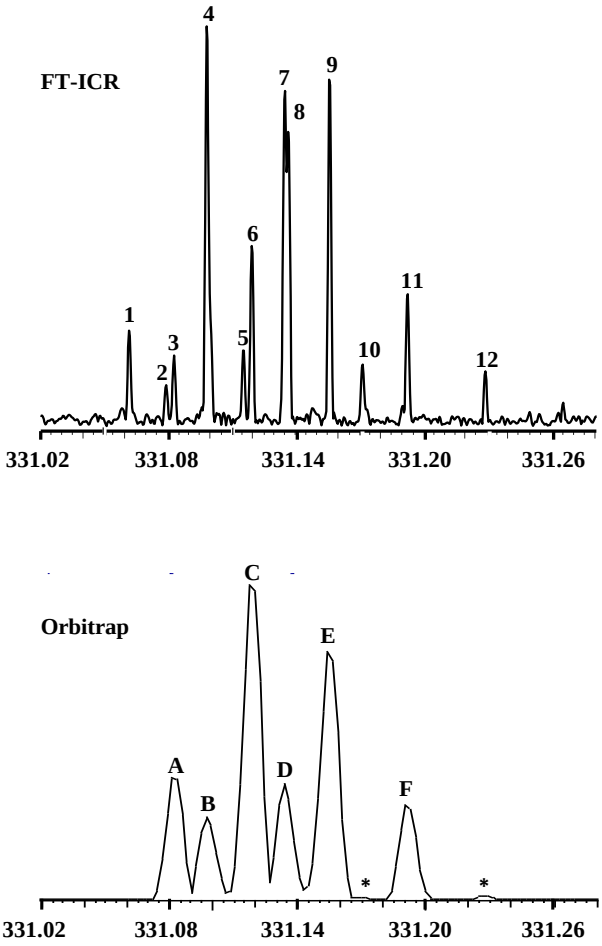
m/z	Formula [M-H]$^-$	Error (ppm)
401.1036	$C_{24}H_{17}O_6$	1.5
401.1245	$C_{21}H_{21}O_8$	0.8
401.1392	$C_{25}H_{21}O_5$	-0.5
401.1606	$C_{22}H_{25}O_7$	0.1
401.1754	$C_{26}H_{25}O_4$	-1.1
401.1971	$C_{23}H_{29}O_6$	0.3

Figure A.2. Expanded regions of Orbitrap mass spectra of PS oily fraction at nominal mass 209 (Panel A) and 401 (Panel B), demonstrating common mass differences observed throughout the spectra of the four bio-oil samples analyzed in this study. Although partial overlapping of peaks was observed (*e.g.*, C_4 *vs* O_3 peaks, corresponding to $\Delta m/z = 0.015$), this overlap did not seem to appreciably effect the *m/z* values of these peaks. That is, chemical formula assignments enabled by Xcalibur peak centroiding algorithms (*e.g.*, table inset in Figure A.2, Panel B) coincided with chemical formulas assigned by propagating the formulas along a homologous series (see text in second paragraph of the Results and Discussion Section).



m/z	Formula $[M-H]^-$	Error (ppm)
431.1498	$C_{26}H_{23}O_6$	-0.6
431.1858	$C_{27}H_{27}O_5$	-1.5
431.2239	$C_{27}H_{27}O_6$	2.6

Figure A.3. Expanded region of Orbitrap mass spectrum of PS oily fraction at nominal mass 431, showing that peaks with $\Delta m/z$ = 0.015 Da were not resolved in some instances and appeared as shouldered peaks (shoulders are shown by arrows). However, the common difference of CH_4 *vs.*O (such as those above, *e.g.*, exact $\Delta m/z$ = 0.0364) was well resolved throughout the mass spectra of all bio-oils analyzed in this study.



Peak	m/z	Formula [M-H]$^-$	Error (ppm)
1	331.0614	$C_{20}H_{11}O_5$	0.6
2	331.0786	$C_{12}H_{15}O_9N_2$	0.9
3	331.0824	$C_{17}H_{15}O_7$	0.2
4	331.0978	$C_{21}H_{15}O_4$	0.7
5	331.1149	$C_{13}H_{19}O_8N_2$	0.6
6	331.1188	$C_{18}H_{19}O_6$	0.3
7	331.1344	$C_{22}H_{19}O_3$	0.4
8	331.1357	$C_{10}H_{23}O_{10}N_2$	1.3
9	331.1552	$C_{19}H_{23}O_5$	0.3
10	331.1705	$C_{23}H_{23}O_2$	0.4
11	331.1916	$C_{20}H_{27}O_4$	0.3
12	331.2280	$C_{21}H_{31}O_3$	0.4

Peak	m/z	Formula [M-H]$^-$	Error (ppm)
A	331.0823	$C_{17}H_{15}O_7$	0.0
B	331.0974	$C_{21}H_{15}O_4$	0.6
C	331.1186	$C_{18}H_{19}O_6$	0.3
D	331.1337	$C_{22}H_{19}O_3$	0.9
E	331.1548	$C_{19}H_{23}O_5$	0.9
F	331.1913	$C_{20}H_{27}O_4$	0.6

Figure A.4. Expanded regions of representative (−)ESI FT–ICR (upper panel) and (−)ESI–Orbitrap (lower panel) mass spectra of the PS oily fraction at nominal mass 331. The FT–ICR mass spectrum shows six additional peaks compared to peaks observed in the Orbitrap mass spectrum. Although peak 1 is clearly absent in the Orbitrap mass spectrum, it appears that peaks 10 and 12 may be present at relatively low intensity (indicated by an asterisk in the Orbitrap mass spectrum). The observed difference in presence and/or relative abundance of ions in Orbitrap versus FT-ICR mass spectra is not unexpected because of the different instruments and instrumental conditions. For the apparent unresolved peaks 2-3, 5-6, and 7-8, it is not possible to be certain whether the extra peaks observed in the FT-ICR mass spectrum are present but unresolved in the Orbitrap mass spectrum or are actually absent as observed for peak 1. However, the Orbitrap data accurately identified the same chemical formula as one of the identified components in the FT-ICR mass spectrum in each case, suggesting that if the extra peak was present, the observed mass was still sufficiently close to that of the more abundant peak in the FT-ICR mass spectrum to identify that molecular formula. It is also important to realize that our approach (*i.e.*, propagating molecular formulas along a homologous series) would support accurate assignments even if an unresolved peak was responsible for a small shift in measured mass

FT-ICR Data Acquisition

The FT-ICR spectrum shown in Figure A.3 was acquired using an IonSpec FT-ICR MS system (former IonSpec Corp. - now a division of Agilent Technologies, Santa Clara, CA) equipped with an open-ended cylindrical ICR cell and a 9.4 tesla superconducting magnet (Cryomagnetics Inc., Oakridge, TN). Ions were generated externally using an Analytica ESI source (Analytica of Branford Inc., Branford, CT) equipped with an in-house built nano-spraying setup. The flow rate was set at 0.5 uL/min using a Harvard PHD 2000 syringe pump (Harvard Apparatus, Holliston, MA). ESI voltage was set to –3 kV. ICR cell pressure was measured by direct reading of a Granville-Phillips dual ion gauge controller and series 274 Bayard-Alpert type ionization gauge tube. Typical base pressure in the ICR cell region was ~3.6×10^{-10} Torr (not corrected for geometry factor, ionization sensitivity, and magnetic field effect,). The trapped ions were excited (for 4 *ms*) using dipolar frequency sweep excitation [M. B. Comisarow and A. G. Marshall, *Chem. Phys. Lett.*, 1974, **26**, 489-490] and detected in broadband mode. Fourier transformation of the acquired 10 time-domain signals (2048 k data points) with one zero fill and Blackman window apodization, followed by magnitude calculation and frequency-to-*m/z* conversion, yielded the ESI/FT-ICR mass spectrum shown in Figure A.3.

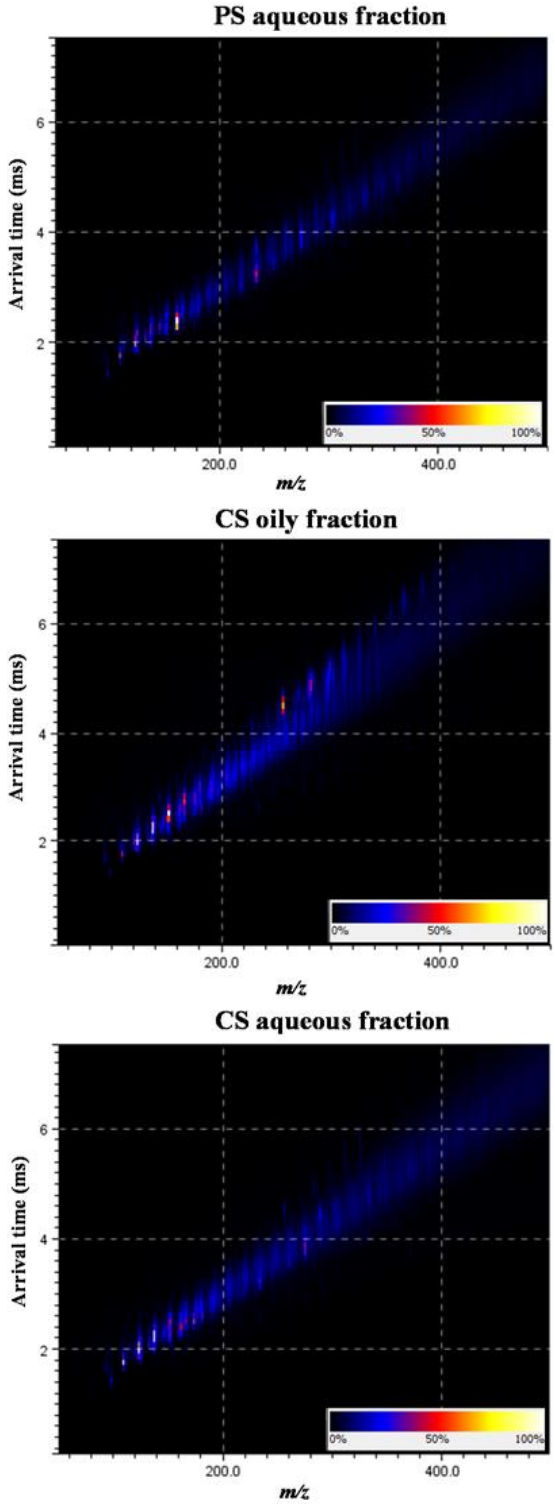


Figure A.5. Two-dimensional (–)ESI-IM-MS spectra showing the full MW distribution of the pine shavings (PS) aqueous fraction (top panel), corn stover (CS) oily fraction (middle panel) and CS aqueous fraction (bottom panel). Note that comparative data for the PS oily fraction is shown in Figure 2.5A.

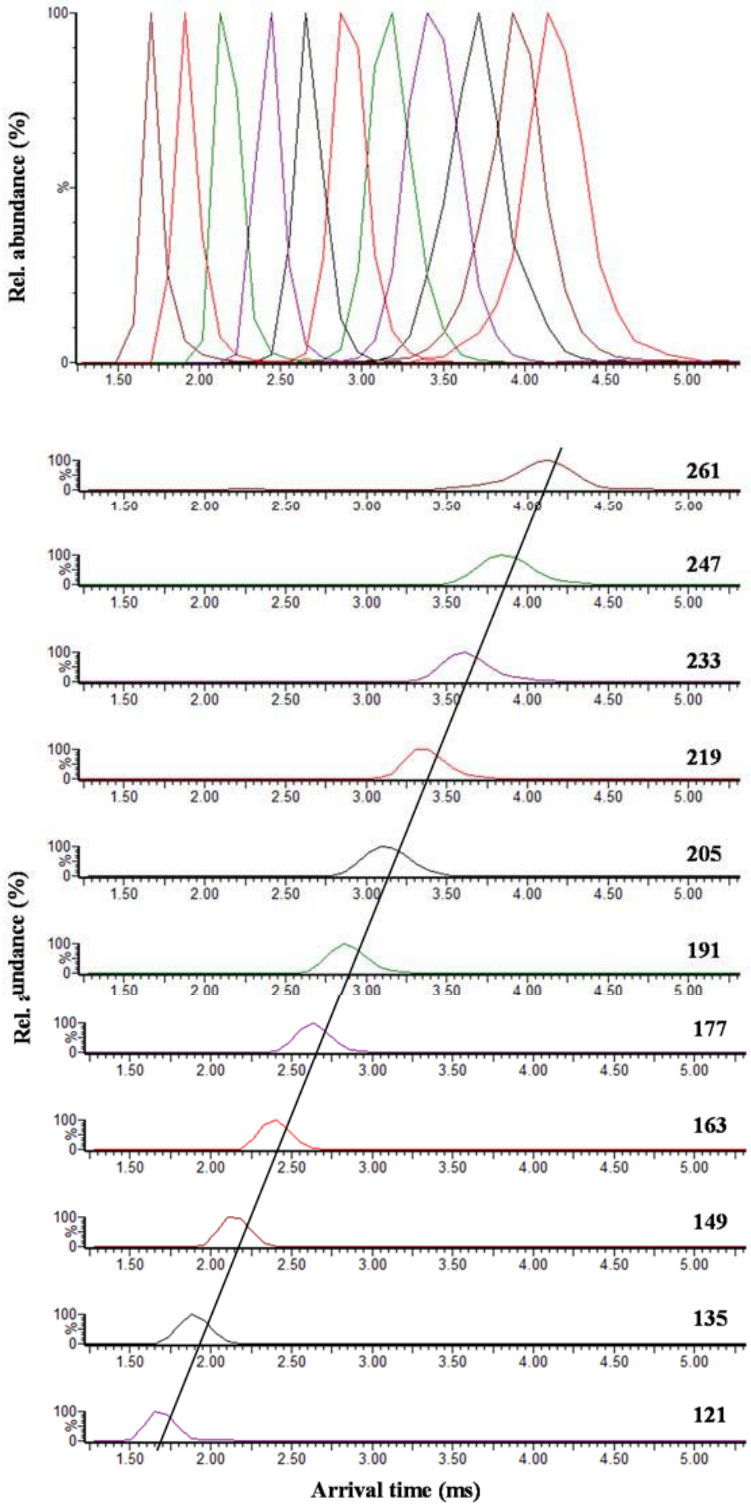


Figure A.6. Representative data showing ion mobility arrival time distributions (ATD) for $C_nH_{2n-8}O_2$ (*i.e.*, series 1 in Figure 2.7). The analyte *m/z* is indicated by the number on the right side of each individual ATD. Each ATD is obtained by integrating across the corresponding to the monoisotopic peak of the analyte.

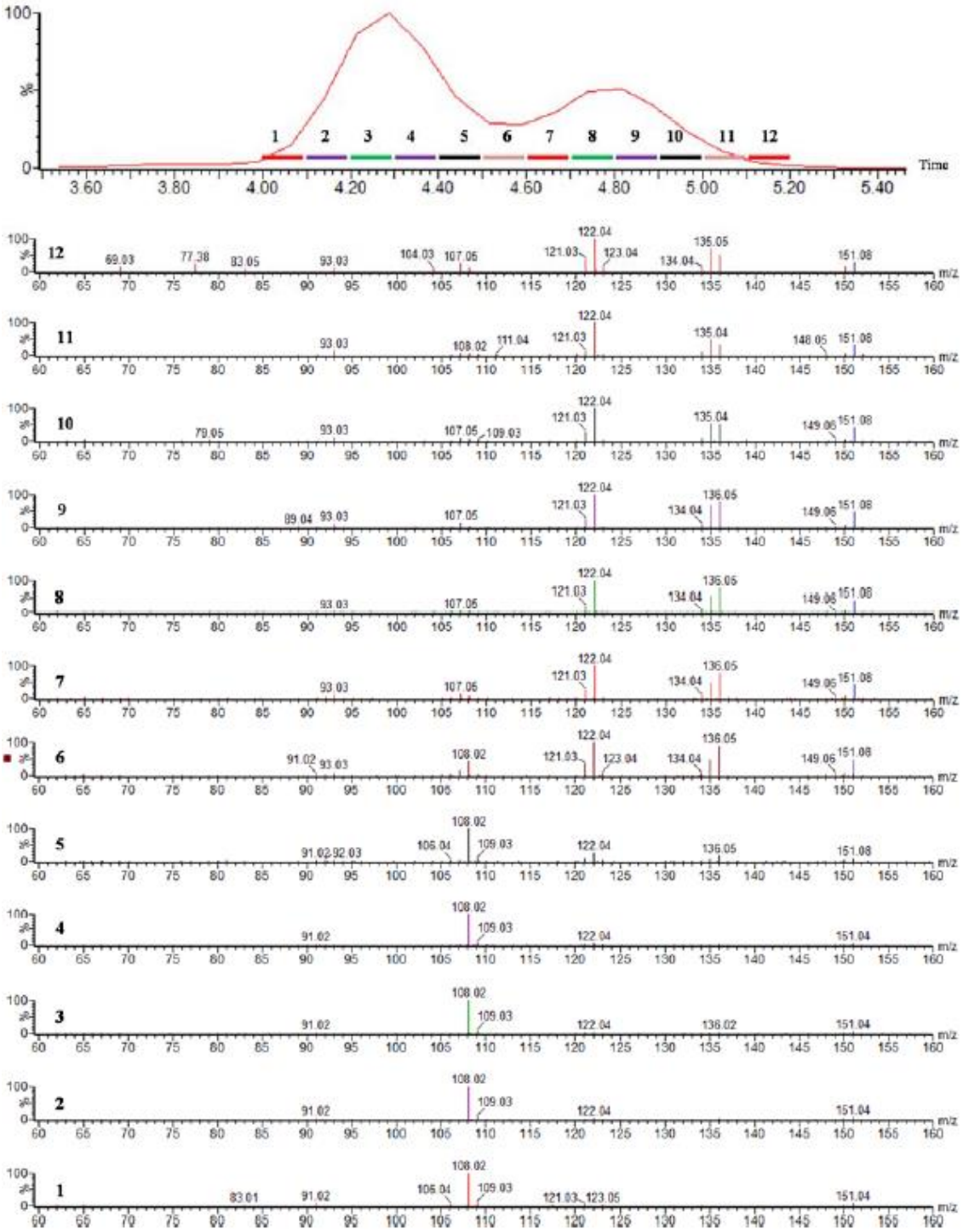


Figure A.7. CID spectra of mobility separated peaks are shown as a function of arrival time. Mass spectra 1-12 were obtained by sampling 0.1 ms sections across each mobility distribution (shown at the top of the figure). The ions shown correspond to nominal mass 151 and are identical to those pictured in Figure 2.8. Note that sampling smaller mobility regions did not provide additional information, when compared to CID spectra shown in Figure 2.9.

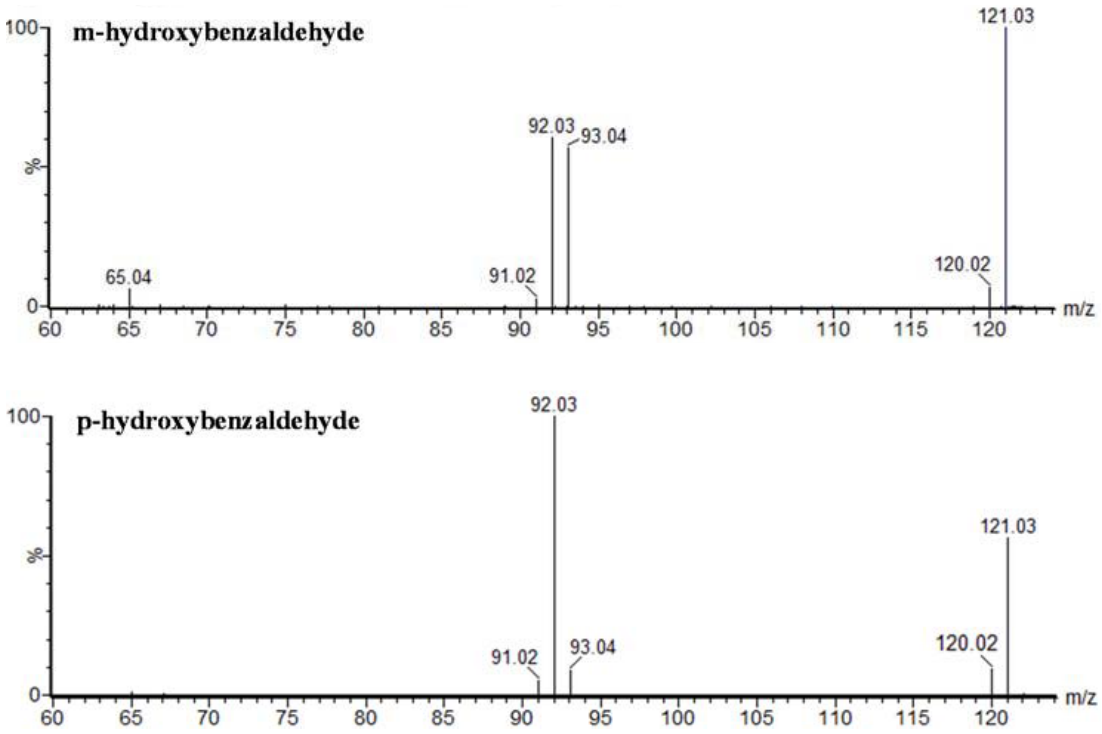


Figure A.8. CID spectra of commercial standards m-hydroxybenzaldehyde (top) and phydroxybenzaldehyde (bottom). Similar fragmentation pattern of these compounds were also reported by Attygalle *et al*. [Attygalle, A. B.; Ruzicka, J.; Varughese, D.; Bialecki, J. B.; Jafri, S.J. *Mass Spectrom*. 2007,*42*, 1207−1217]. CID spectrum of the precursor ion shown in Figure 2.10A (*i.e*., *m/z* 121.0295) matches well with spectrum for *m*-hydroxybenzaldehyde. Note that it is not clear whether the difference in the relative abundances of *m/z* 92 and 93 in Figure 2.10A is due to the experimental uncertainty inherent when comparing relative abundances, or if there is minor contribution from the presence of *p*-hydroxybenzaldehyde.

Table A.1. Average RMS errors of chemical formula assignments for all identified analytes in each fraction

Bio-oil sample	Average error (ppm)
PS oily fraction	0.9
PS aqueous fraction	1.4
CS oily fraction	1.0
CS aqueous fraction	1.3

Molecular Formula Assignments in Table A.2

We are presenting data in the given tabular format because it is concise and also because it facilitates comparison among different samples. The table organizes all compounds according to the number of oxygen atoms in their molecular formula (*e.g.*, their oxygen class). The second column lists the double bond equivalence (DBE) for each compound in that row while the remaining columns show the number of carbons for each chemical formula identified in a given feedstock having a specific oxygen class and DBE value. In this manner a single row can represent numerous chemical formulas. From the information given in Table A.2, the molecular formula $C_cH_hO_oN_n$, can be obtained where "*c*", "*o*", and "*n*" are retrieved from their respective columns and the number of hydrogen atoms "*h"* is obtained using the following equation:

$h = 2c + 2 + n - 2$*DBE

For example, in the PS oily fraction (PSOF) the O_3-class row with a DBE of 4 identifies twelve compounds, *i.e.*, compounds containing 3 oxygen atoms and 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, and 18 carbons. Molecular formulas corresponding to each of these compounds can be calculated using the above equation such that we obtain the following molecular formulas: $C_5H_4O_3$, $C_6H_6O_3$, $C_7H_8O_3$ … $C_{15}H_{24}O_3$, and $C_{18}H_{30}O_3$.

Table A.2. Heteroatom class, DBE and number of carbons of compounds identified in oily fraction (OF) and aqueous fraction (AF) of pine shavings (PS) and corn stover (CS)

		Number of carbons			
Class	**DBE**	**PSPOF**	**PSAF**	**CSOF**	**CSAF**
O_1	1	6, 8, 9		6, 8	
O_1	2	5–9	4–8	5–9	4–9
O_1	3	5–10	5–9	5–10	5–9, 14
O_1	4	6–12	6–10	6–12	6–10
O_1	5	6–13	6–11	6–13	6–10
O_1	6	8–14	8–11	8-14	8–10
O_1	7	9–15		10–15	11
O_1	8	12–15		12–16	
O_1	9	15		14–17	
O_1	10	15–18		14–18	
O_1	12	18, 19		11, 13, 15	
O_1	14	21, 22			
O_2	1	4–10, 14–22, 24, 26	2–7, 16, 18	4–10, 12, 14–26	2–6, 8, 11, 19, 20
O_2	2	4–10, 16–20	3–9	4–11, 14–20, 22, 24	4–8, 11
O_2	3	5–14, 18	4–10	5–14, 16, 18,	5-14
O_2	4	5–15,18	5–11,14	5–15,18	5–11, 13–15
O_2	5	7–15, 18	6–12	7–16, 18	7–12
O_2	6	8–18, 20	8–13	8–20	8–10, 12, 13
O_2	7	9–20		9, 10, 12–20	9, 13
O_2	8	11–20		11–20	25
O_2	9	12–17, 19, 20		12–17, 19, 20	
O_2	10	13–21		13–20	
O_2	11	17–21		17–21	
O_2	12	17–21		17–21	
O_2	13	20, 21			
O_2	14	20, 21			
O_2	15	23			
O_3	1	5, 7–10, 16, 18	2–5	4, 5, 7, 9, 10, 14, 16, 18–26	3–5
O_3	2	5–10, 16–18	2–7	4–10, 16, 18	3–9, 11
O_3	3	4–13, 18	4–9	4–12, 18	4–10
O_3	4	5–15,18	5–12	5–15, 18	5–10, 12, 14

Table A.2. Heteroatom class, DBE and number of carbons of compounds identified in oily fraction (OF) and aqueous fraction (AF) of pine shavings (PS) and corn stover (CS)

		Number of carbons			
Class	**DBE**	**PSPOF**	**PSAF**	**CSOF**	**CSAF**
O_3	5	6–18	6–13	6–18	6–13
O_3	6	7–20	7–14	8–18	8–11, 13, 14
O_3	7	9–20	9,12–15	9–20	9, 10, 13–15
O_3	8	10–21	10–14, 16,17	10–21	15
O_3	9	12–20	12–16	12–18	13–15
O_3	10	13–22	14, 15	13–22	
O_3	11	15–23	15	15–23	
O_3	12	16–24		17–23	
O_3	13	18–24		19–23	
O_3	14	21–25		22, 23	
O_3	15	23, 24			
O_4	1	5, 18	4–6	4–6, 18, 24	4–6
O_4	2	5–11,18	3–9	5–10,18	4–9
O_4	3	4–11, 18	5–10	4–11, 18	5–9
O_4	4	6–14,18	6–12	7–14	8–11, 13
O_4	5	7–14, 16–18	6–12, 17	7–14, 16–18	8–12, 16
O_4	6	8–20	7–14	7–19	7–14
O_4	7	9–20	8–12, 14–16	9–20	10–12, 14–16
O_4	8	10–21	10–14, 16, 17	10–21	12–16
O_4	9	12–22	11–17	12–22	13–16
O_4	10	13–24	13–18	13–23	
O_4	11	15–24	17, 18	15–24	
O_4	12	16–25	19, 20	17–24	
O_4	13	18–25		19–24	
O_4	14	20–26		21–25	
O_4	15	22–25			
O_5	1	6, 8	5, 6, 8	5, 6	5, 6, 8
O_5	2	6, 7, 9	5–10	6–9	5–9
O_5	3	6, 7, 9	6–10	7–9	7, 9, 10
O_5	4	8–13	7–11	8–13	11, 16
O_5	5	8–13, 15	8–13	8–13, 15	10–12, 16–17
O_5	6	9–19	9–13, 15	9–18	10–14
O_5	7	10–20	8, 10–17	10–20	11–15

Table A.2. Heteroatom class, DBE and number of carbons of compounds identified in oily fraction (OF) and aqueous fraction (AF) of pine shavings (PS) and corn stover (CS)

		Number of carbons			
Class	**DBE**	**PSPOF**	**PSAF**	**CSOF**	**CSAF**
O_5	8	11–21	10–13	12–21	12–14
O_5	9	12–23	11–14, 17–19	13–22	14–18
O_5	10	13–24	12–20	14–23	17
O_5	11	15–25	15–20	16–23	
O_5	12	17–25		18–24	
O_5	13	19–26		20–25	
O_5	14	20–27		22–25	
O_5	15	23–27			
O_6	1		6	5	6
O_6	2	10	6–11	7, 8	7, 8
O_6	3	8	7–11		
O_6	4	11	9–12	10, 11	
O_6	5		10–14	11–13	
O_6	6	11–17	9–15	11–17	
O_6	7	12–19	11–17	12–19	13–15
O_6	8	12–21	11–14,	13–21	14, 16
O_6	9	13–22	11–14	14–22	
O_6	10	14–24	12–21	15–23	17, 18
O_6	11	16–24	15–21	17–24	
O_6	12	17–25	17–21	19–23	
O_6	13	19–26	19, 20	21–25	
O_6	14	21–27			
O_6	15	24–27			
O_7	2		7–11		8
O_7	3		8–11		9
O_7	4		10, 11		
O_7	5		11–13		
O_7	6		12–15		
O_7	7	13–17	12–17	14–17	
O_7	8	14–19	12–19	15–19	10
O_7	9	15–21	12–15	17–20	19
O_7	10	15–23	18–20	17–22	18–20
O_7	11	16–24	15–22	19–22	

Table A.2. Heteroatom class, DBE and number of carbons of compounds identified in oily fraction (OF) and aqueous fraction (AF) of pine shavings (PS) and corn stover (CS)

		Number of carbons			
Class	**DBE**	**PSPOF**	**PSAF**	**CSOF**	**CSAF**
O_7	12	19–25	18–21		
O_7	13	21–25			
O_7	14	22–26			
O_8	4		11		
O_8	5		13		
O_8	6		13–15		
O_8	7		13–16		
O_8	8		14–18		
O_8	8		15, 19		
O_8	10	18–20	14–17		
O_8	11	21, 22	17, 18, 20		
O_8	12	22	19–22		
O_8	13	22-24			
O_1N_2	3				8
O_1N_2	4				7, 8
O_1N_4	13		15–20		16–19
O_1N_4	14		14, 16, 17		
O_2N_4	10		16–21		
O_3N_2	7				13, 14
O_3N_2	8				13, 14
O_3N_2	13		18		
O_3N_4	2		4		
O_3N_4	14		17–22		
O_3N_4	15		18		
O_4N_4	14		18–22		
O_4N_4	15		19,20, 22		

Table A.3. Fragment ions and their relative abundances of several precursor ions

Precursor ion	**Fragment ions (*m/z*, Da)**																	
Sample	Relative Abundance (%)																	
109.03	**109.03**	**108.02**	**91.02**	**65.00**														
PSOF	100	90	15	7														
PSAF	100	94	12	5														
CSOF		100	15	5														
CSAF		100	15	7														
123.04	**122.04**	**121.03**	**108.02**	**105.03**														
PSOF	100	14	32	9														
PSAF	100	10	23	5														
CSOF	100	12	24	6														
CSAF	100	13	28	7														
137.02	**136.02**	**119.01**	**109.03**	**108.02**	**93.03**	**92.03**	**91.02**	**65.00**										
PSOF	77	7	21	100	10	38	26	11										
PSAF	80	6	23	100	9	21	19	5										
CSOF	58	6	16	100	16	26	16	6										
CSAF	64	6	22	100	11	28	21	5										

Table A.3. Fragment ions and their relative abundances of several precursor ions – Continued

Precursor ion	**Fragment ions (*m/z*, Da)**																	
Sample	Relative Abundance (%)																	
137.06	**136.05**	**135.05**	**122.04**	**121.03**	**107.05**	**93.03**												
PSOF	80	19	100	17	11	11												
PSAF	94	15	100	19	6	< 5												
CSOF	56	9	100	21	< 5	< 5												
CSAF	49	11	100	24	< 5	6												
151.04	**122.04**	**109.03**	**108.02**	**91.02**														
PSOF	< 5	11	100	< 5														
PSAF	< 5	8	100	< 5														
CSOF	9	11	100	< 5														
CSAF	5	12	100	7														
151.08	**150.07**	**149.06**	**136.05**	**135.05**	**134.04**	**122.04**	**121.03**	**108.02**	**107.05**	**93.03**								
PSOF	< 5	< 5	98	41	< 5	100	16	13	7	6								
PSAF	< 5	24	100	58	8	68	18	4	13	7								
CSOF	9	< 5	100	57	7	73	22	2	11	7								
CSAF	10	11	100	99	17	78	26	8	21	15								

Table A.3. Fragment ions and their relative abundances of several precursor ions – Continued

Precursor ion	**Fragment ions (*m/z* , Da)**																		
Sample	Relative Abundance (%)																		
165.06	**150.03**	**123.04**	**122.04**	**121.03**	**109.03**	**108.02**	**105.04**	**93.03**											
PSOF	5	12	83	6	5	100	5	< 5											
PSAF	7	23	94	11	5	100	5	6											
CSOF	7	13	100	12	6	82	7	< 5											
CSAF	5	16	100	9	5	54	5	< 5											
165.00	**150.07**	**149.06**	**148.05**	**136.00**	**136.05**	**135.05**	**134.04**	**122.04**	**121.03**	**93.03**									
PSOF	16	39	5	41	100	9	25	7	31	5									
PSAF	26	53	15	54	100	13	35	19	49	14									
CSOF	18	37	7	46	100	9	46	14	39	5									
CSAF	31	44	11	27	100	30	37	17	33	11									
121.07	**121.07**	**120.06**	**119.05**	**106.04**															
PSOF	100	8	< 5	55															
CSOF	100	8	7	53															
121.03	**120.02**	**93.03**	**92.03**	**91.02**	**65.04**														
PSOF	8	61	100	5	9														
CSOF	9	17	100	5	3														

Table A.3. Fragment ions and their relative abundances of several precursor ions – Continued

Precursor ion	**Fragment ions (*m/z*, Da)**																	
Sample	Relative Abundance (%)																	
135.04	**134.04**	**120.02**	**107.05**	**106.04**	**93.04**	**92.03**	**89.04**											
PSOF	64	6	15	34	22	100	12											
PSAF	29	6	7	15	15	100	8											
CSOF	27	6	8	15	21	100	< 5											
CSAF	9	5	5	5	17	100	< 5											
135.08	**135.08**	**133.07**	**120.58**	**119.05**	**106.04**	**92.03**												
PSOF	100	10	93	53	42	22												
CSOF	62	8	100	49	34	17												
149.06	**148.05**	**147.05**	**146.37**	**134.04**	**133.03**	**129.04**	**122.04**	**121.03**	**108.02**	**107.05**	**106.04**	**101.04**	**93.04**	**92.03**	**65.04**			
PSOF	37	31	< 5	14	20	13	11	6	25	21	100	9	9	33	< 5			
PSAF	39	47	7	13	17	27	< 5	7	17	15	100	20	6	31	6			
CSOF	23	22	< 5	16	9	14	8	< 5	11	18	100	9	5	36	< 5			
CSAF	17	12	< 5	< 5	7	5	< 5	< 5	5	16	100	7	5	27	5			
163.08	**162.69**	**161.06**	**148.05**	**147.05**	**146.04**	**133.03**	**129.04**	**122.04**	**121.07**	**120.06**	**119.05**	**108.02**	**107.05**	**106.04**	**101.04**	**92.03**	**77.04**	
PSOF	9	14	25	42	12	23	14	12	13	53	9	10	11	100	10	23	< 5	
PSAF	18	37	35	85	43	18	20	13	29	64	11	13	18	100	16	28	15	
CSOF	7	15	29	48	12	18	12	15	14	54	12	10	13	100	11	26	< 5	
CSAF	8	14	46	37	16	21	10	8	26	74	24	7	15	100	9	23	26	

Table A.3. Fragment ions and their relative abundances of several precursor ions – Continued

Precursor ion	**Fragment ions (*m/z*, Da)**																	
Sample	Relative Abundance (%)																	
177.09	**162.07**	**161.06**	**159.05**	**148.05**	**147.05**	**146.04**	**135.05**	**134.04**	**133.03**	**129.04**	**122.04**	**121.07**	**120.06**	**119.52**	**117.03**	**106.04**	**105.04**	**92.03**
PSOF	28	96	19	19	82	48	20	21	17	21	24	19	72	23		100		29
CSOF	35	94	21	31	95	34	26	15	18	20	18	14	71	25	19	100	21	27

APPENDIX B

The Influence of Terminal Group and Repeat Unit Structure on Mass – Mobility Correlations Observed for Homologous Series

Data and the affiliated discussion reported in this section summarize the approach taken to evaluate potential clusters formed by methanol and target analytes in the present study.

Figure B.1A displays the mass spectrum observed for Series 8 (*i.e.*, the amino-carboxylic acid series ionized in positive-ion mode ESI). These data clearly support that peaks were not observed at *m/z* values corresponding to adducts formed by protonated molecular ions and methanol. Although the small peak at *m/z* 182 has the correct mass for a cluster involving the protonated molecular ion at *m/z* 118 (n = 4) and two methanol molecules, we believe this is merely coincidental, since peaks for analogous clusters involving two methanol molecules and $[M + H]^+$ ions at *m/z* 132, 160 and 216 (n = 5, 7, and 11, respectively) were not observed. Figure B.1B displays the mass spectrum observed for Series 2 (*i.e.*, the aniline series). This series exhibited the most positive slope in Figure 3.2. However, by similar argument to that given above for Series 8, these data also show no evidence supporting formation of adducts or clusters with methanol.

Previous studies focused on understanding energy conditions along the ion path in Synapt instruments have demonstrated that, under normal conditions, ion fragmentation is more likely to occur before or in the mobility region of the instrument than it is further downstream (Merenbloom et al. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 553-562 and Morsa et al. *Anal. Chem.* **2011**, *83*, 5775–5782). Thus, it is unlikely that ion clusters

contributed to measured arrival times and were completely declustered between the mobility cell and the detector in the present study. That said, mass spectra in Figure B.1 are technically inconclusive with respect to such a possibility. An increase in arrival time would obviously be expected for the larger (*i.e.*, clustered) ions. One might also expect to see a change in peak width and/or peak asymmetry. The inset in Figure B.1A shows

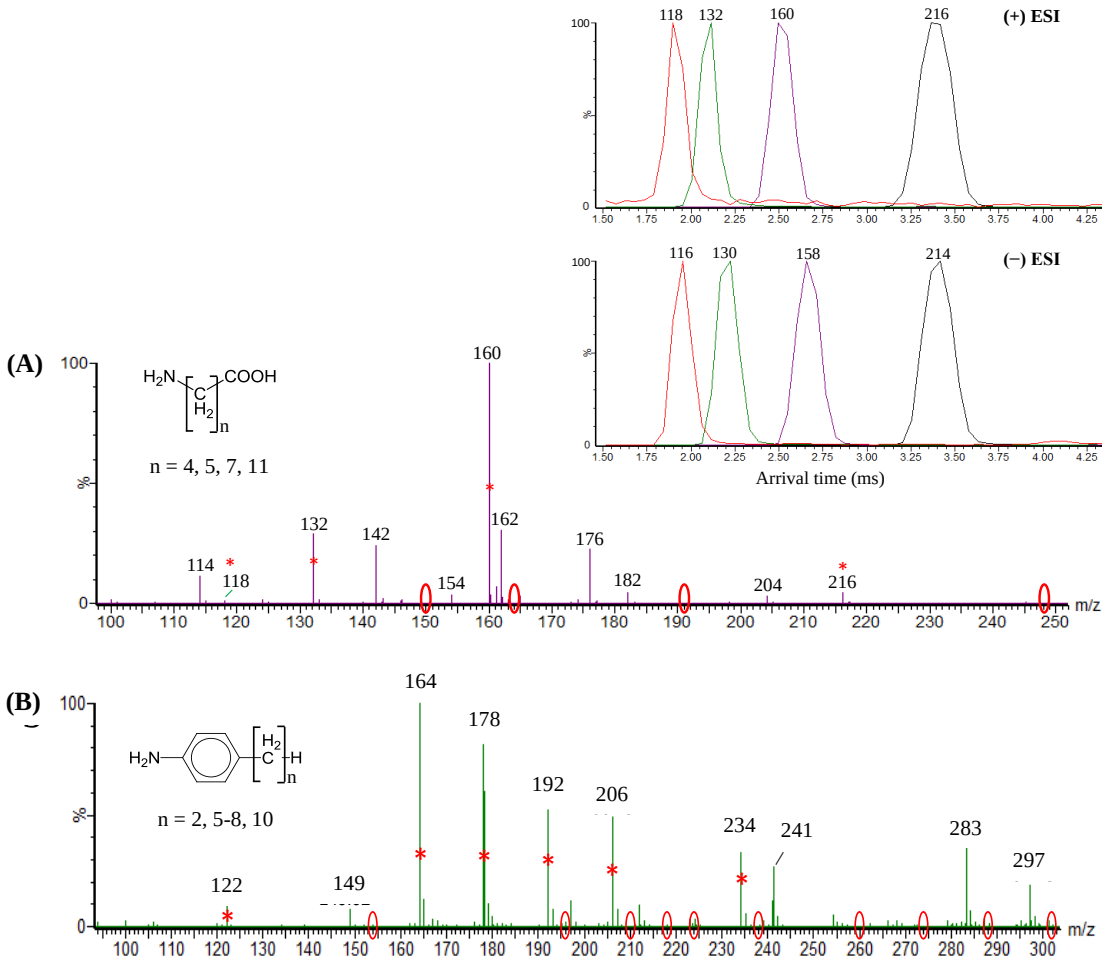


Figure B.1 Representative (+)ESI-MS spectra observed for (A) Series 8 and (B) Series 2. Peaks designated by asterisks correspond to the protonated molecular ions (i.e., [M + H]+) observed for each investigated member of these CH_2-homologous series. Ellipsoids designate *m/z* values where adducts or clusters formed between $[M + H]^+$ ions and methanol would be expected to appear. The inset in Figure B.1A shows arrival time distributions observed for the amino-carboxylic acid series when members were ionized in positive and negative ESI modes.

ATDs observed for Series 7 and 8 (*i.e.*, the amino-carboxylic acid series ionized in negative and positive ESI modes, respectively). These data clearly demonstrate that significant differences in arrival times and peak shapes were not observed for corresponding $[M + H]^+$ and $[M – H]^–$ ions.

An independent experiment was performed to evaluate whether adducts/clusters were contributing to arrival times measured for Series 2 (*i.e.*, the aniline series). In this case, increasing voltage bias was applied to the trap cell preceding the mobility region of the instrument. If clusters were present and contributing to measured arrival times, their effect would be most prevalent at low trap voltage (*i.e.*, 2V corresponding to the normal condition used to collect all data reported in the present manuscript). Higher trap voltages would be expected to promote declustering prior to the mobility separation, and this would result in a decrease in arrival times observed at higher energy. As shown in Figure B.2, arrival times were essentially independent of trap voltage even when the energy was sufficiently high to initiate fragmentation of $[M + H]^+$ ions (*i.e.*, 20 or 30 V). The argument favoring fragmentation at higher trap voltages was also supported by an observed decrease in absolute intensity for *m/z* 192. Detector counts for this peak were reduced from 1.95×10^6 at a trap voltage 2V to 7.51×10^5 and 3.06×10^4 at trap voltages of 20 and 30 V, respectively. These data clearly suggest that methanol clusters did not contribute to measured arrival times for members of this series.

A final experiment was performed to evaluate the effect of changing the spray solvent. Figure B.3 shows the resulting mass spectrum when two members of the aniline series were ionized from a 95% (v/v) acetonitrile-water solution. Acetonitrile is aprotic and would not be expected to cluster in the same way that methanol does. There is no

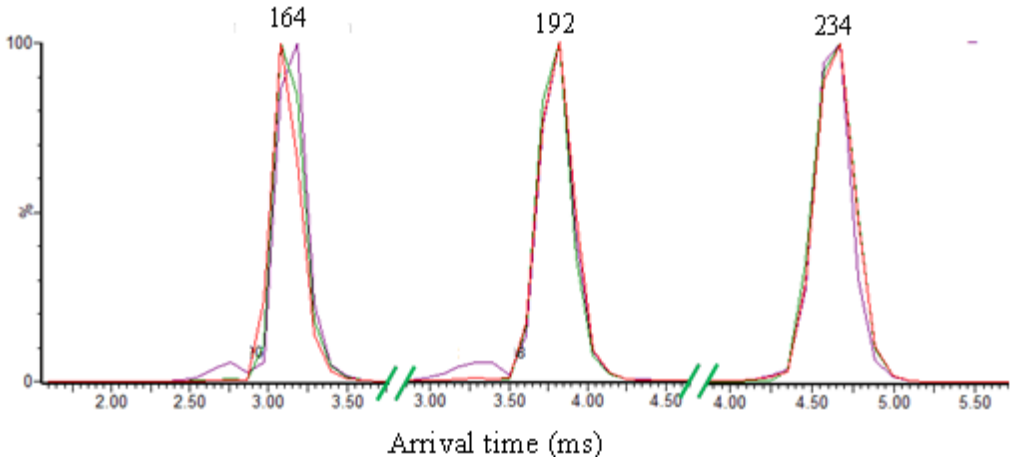


Figure B.2. Arrival time distributions observed for representative members of Series 2 at various trap voltages: 2V (red), 20 V (green), 30 V (purple). Peaks labels correspond to *m/z* values appearing in Figure B.1B.

evidence for adduct/cluster formation in the mass spectrum. More importantly, the ATDs observed for these ions were virtually identical when either acetonitrile or methanol was the primary component of the spray solvent. Clusters formed with different solvents would be expected to produce substantially different ATDs. Thus, this result also suggests an absence of clustering in both solvents.

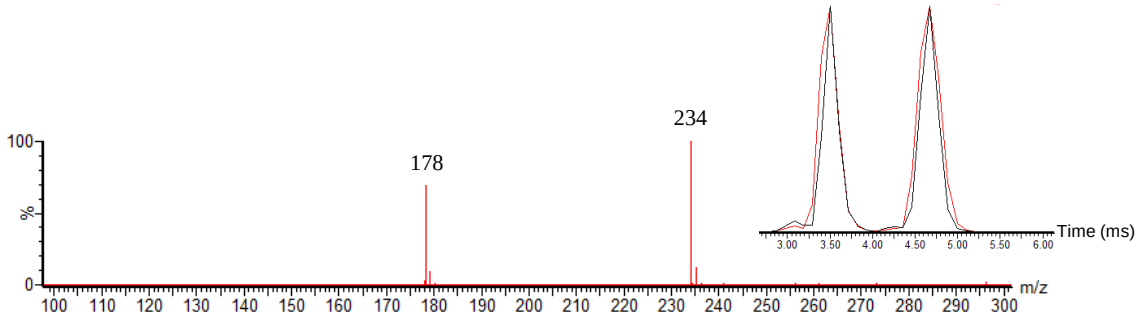


Figure B.3 Mass spectrum observed when compounds from Series 2 with n = 6 and 10 were ionized using an alternative spray solvent. The inset shows corresponding ATDs for these compounds when 95% acetonitrile (red) and methanol (black) were used as the spray solvent.

APPENDIX C

Characterization of Various Solvent Fractions of Slow-pyrolysis Bio-oil by High-resolution Mass Spectrometry and Ion Mobility Spectrometry

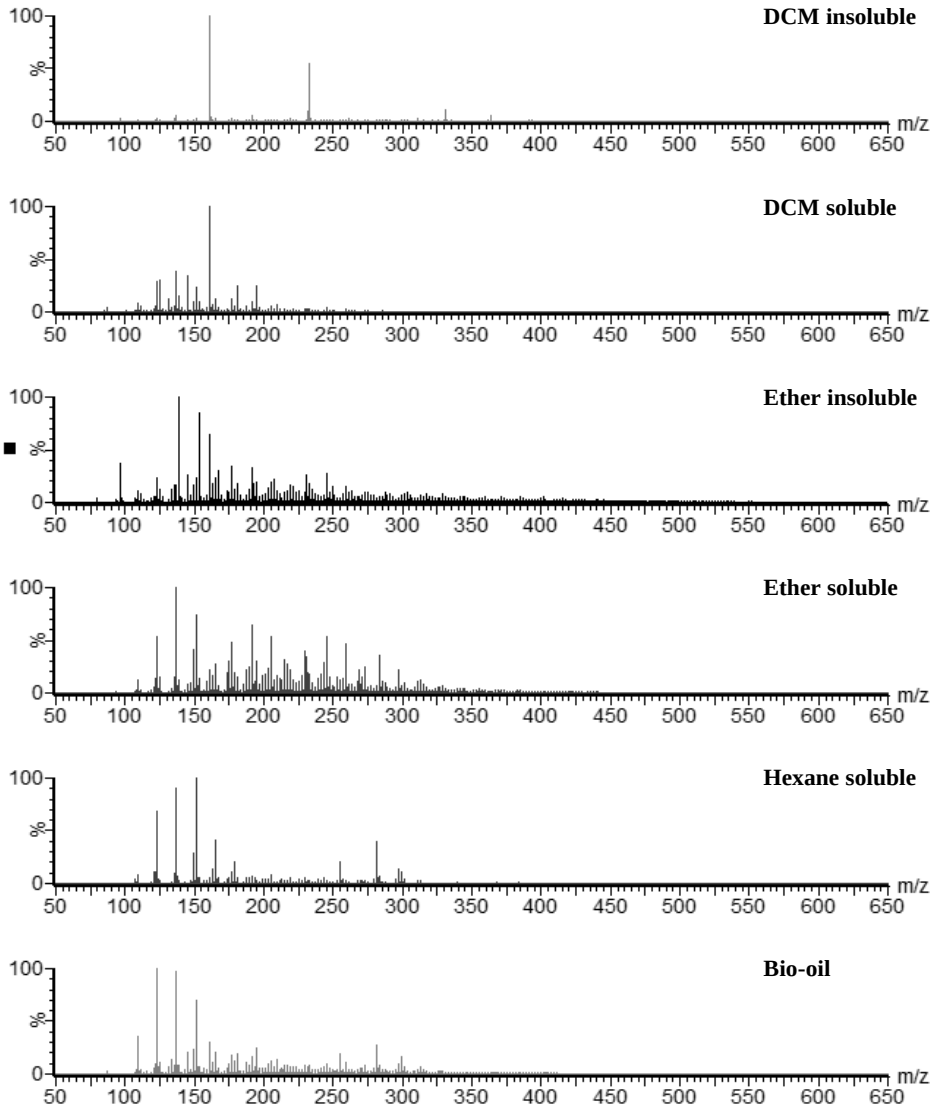


Figure C.1. Molecular weight distibutions of bio-oil and its solvent fractions observed in (-)ESI-TOF-MS (Synapt G2-S) analyses.

REFERENCES

(1) Höök, M.; Tang, X. *Energy Policy* **2013**, *52*, 797–809.

(2) Guo, M.; Song, W.; Buhain, J. *Renew. Sustain. Energy Rev.* **2015**, *42*, 712–725.

(3) Capunitan, J. A.; Capareda, S. C. *Fuel* **2012**, *95*, 563–572.

(4) Xin, S.; Yang, H.; Chen, Y.; Wang, X.; Chen, H. *Fuel* **2013**, *113*, 266–273.

(5) Bright, R. M.; Strømman, A. H. *J. Ind. Ecol.* **2009**, *13* (4), 514–531.

(6) Akhtar, J.; Saidina Amin, N. *Renew. Sustain. Energy Rev.* **2012**, *16* (7), 5101–5109.

(7) Zhou, C.-H.; Xia, X.; Lin, C.-X.; Tong, D.-S.; Beltramini, J. *Chem. Soc. Rev.* **2011**, *40* (11), 5588–5617.

(8) Isahak, W. N. R. W.; Hisham, M. W. M.; Yarmo, M. A.; Yun Hin, T. *Renew. Sustain. Energy Rev.* **2012**, *16* (8), 5910–5923.

(9) Oasmaa, A.; van de Beld, B.; Saari, P.; Elliott, D. C.; Solantausta, Y. *Energy Fuels* **2015**, *29* (4), 2471–2484.

(10) Bridgwater, A. V.; Meier, D.; Radlein, D. *Org. Geochem.* **1999**, *30* (12), 1479–1493.

(11) Bridgwater, A. V. *Biomass Bioenergy* **2012**, *38*, 68–94.

(12) Fagernäs, L.; Kuoppala, E.; Tiilikkala, K.; Oasmaa, A. *Energy Fuels* **2012**, *26* (2), 1275–1283.

(13) Huber, G. W.; Iborra, S.; Corma, A. *Chem. Rev.* **2006**, *106* (9), 4044–4098.

(14) Lehto, J.; Oasmaa, A.; Solantausta, Y.; Kytö, M.; Chiaramonti, D. *Appl. Energy* **2014**, *116*, 178–190.

(15) Maggi, R.; Delmon, B. *Fuel* **1994**, *73* (5), 671–677.

(16) Schneider, J. K.; da Cunha, M. E.; dos Santos, A. L.; Maciel, G. P. S.; Brasil, M. C.; Pinho, A. R.; Mendes, F. L.; Jacques, R. A.; Caramão, E. B. *J. Chromatogr. A* **2014**, *1356*, 236–240.

(17) Das, P.; Sreelatha, T.; Ganesh, A. *Biomass Bioenergy* **2004**, *27* (3), 265–275.

(18) Demiral, İ.; Ayan, E. A. *Bioresour. Technol.* **2011**, *102* (4), 3946–3951.

(19) Aysu, T.; Küçük, M. M. *Energy* **2014**, *64*, 1002–1025.

(20) Açıkalın, K.; Karaca, F.; Bolat, E. *Fuel* **2012**, *95*, 169–177.

(21) Ben, H.; Ragauskas, A. J. *Bioresour. Technol.* **2013**, *147*, 577–584.

(22) Oasmaa, A.; Kuoppala, E.; Ardiyanti, A.; Venderbosch, R. H.; Heeres, H. J. *Energy Fuels* **2010**, *24* (9), 5264–5272.

(23) Choi, Y. S.; Johnston, P. A.; Brown, R. C.; Shanks, B. H.; Lee, K.-H. *J. Anal. Appl. Pyrolysis* **2014**, *110*, 147–154.

(24) Garcia-Perez, M.; Chaala, A.; Pakdel, H.; Kretschmer, D.; Roy, C. *Biomass Bioenergy* **2007**, *31* (4), 222–242.

(25) Fagernäs, L.; Kuoppala, E.; Simell, P. *Energy Fuels* **2012**, *26* (11), 6960–6970.

(26) Villadsen, S. R.; Dithmer, L.; Forsberg, R.; Becker, J.; Rudolf, A.; Iversen, S. B.; Iversen, B. B.; Glasius, M. *Energy Fuels* **2012**, *26* (11), 6988–6998.

(27) Bai, X.; Kim, K. H.; Brown, R. C.; Dalluge, E.; Hutchinson, C.; Lee, Y. J.; Dalluge, D. *Fuel* **2014**, *128*, 170–179.

(28) Marshall, A. G.; Rodgers, R. P. *Acc. Chem. Res.* **2004**, *37* (1), 53–59.

(29) Herzsprung, P.; v. Tümpling, W.; Hertkorn, N.; Harir, M.; Friese, K.; Schmitt-Kopplin, P. *Anal. Chem.* **2015**, *87* (19), 9563–9566.

(30) Trompowsky, P. M.; Benites, V. de M.; Madari, B. E.; Pimenta, A. S.; Hockaday, W. C.; Hatcher, P. G. *Org. Geochem.* **2005**, *36* (11), 1480–1489.

(31) Staš, M.; Kubička, D.; Chudoba, J.; Pospíšil, M. *Energy Fuels* **2014**, *28* (1), 385–402.

(32) Alsbou, E.; Helleur, B. *Energy Fuels* **2013**.

(33) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G.; Qian, K. *Anal. Chem.* **2001**, *73* (19), 4676–4681.

(34) Kendrick, E. *Anal. Chem.* **1963**, *35* (13), 2146–2154.

(35) Roach, P. J.; Laskin, J.; Laskin, A. *Anal. Chem.* **2011**, *83* (12), 4924–4929.

(36) Balogh, M. B.; Stranz, D. *LC-GC Eur.* **2011**, *24* (9), 490–499.

(37) Herzsprung, P.; Hertkorn, N.; von Tümpling, W.; Harir, M.; Friese, K.; Schmitt-Kopplin, P. *Anal. Bioanal. Chem.* **2014**, *406* (30), 7977–7987.

(38) Islam, A.; Cho, Y.-J.; Ahmed, A.; Kim, S.-H. *Mass Spectrom. Lett.* **2012**, *3* (3), 63–67.

(39) Kim, S.; Kramer, R. W.; Hatcher, P. G. *Anal. Chem.* **2003**, *75* (20), 5336–5344.

(40) Kind, T.; Fiehn, O. *BMC Bioinformatics* **2007**, *8* (1), 105.

(41) Marshall, A. G.; Hendrickson, C. L. *Annu. Rev. Anal. Chem.* **2008**, *1* (1), 579–599.

(42) Senko, M. W.; Hendrickson, C. L.; Emmett, M. R.; Shi, S. D.-H.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1997**, *8* (9), 970–976.

(43) Shi, S. D.-H.; Drader, J. J.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1999**, *10* (3), 265–268.

(44) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. *J. Mass Spectrom. JMS* **2005**, *40* (4), 430–443.

(45) Peterman, S. M.; Duczak Jr, N.; Kalgutkar, A. S.; Lame, M. E.; Soglia, J. R. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (3), 363–375.

(46) Wu, Z.; Rodgers, R. P.; Marshall, A. G. *Anal. Chem.* **2004**, *76* (9), 2511–2516.

(47) Cho, Y.; Ahmed, A.; Islam, A.; Kim, S. *Mass Spectrom. Rev.* **2015**, *34* (2), 248–263.

(48) Smith, E. A.; Park, S.; Klein, A. T.; Lee, Y. J. *Energy Fuels* **2012**, *26* (6), 3796–3802.

(49) Jarvis, J. M.; McKenna, A. M.; Hilten, R. N.; Das, K. C.; Rodgers, R. P.; Marshall, A. G. *Energy Fuels* **2012**, *26* (6), 3810–3815.

(50) Liu, Y.; Shi, Q.; Zhang, Y.; He, Y.; Chung, K. H.; Zhao, S.; Xu, C. *Energy Fuels* **2012**, *26* (7), 4532–4539.

(51) Smith, E. A.; Lee, Y. J. *Energy Fuels* **2010**, *24* (9), 5190–5198.

(52) Croce, A.; Battistel, E.; Chiaberge, S.; Spera, S.; Reale, S.; De Angelis, F. *Energy Fuels* **2015**, *29* (9), 5847–5856.

(53) Chiaberge, S.; Leonardis, I.; Fiorani, T.; Cesti, P.; Reale, S.; Angelis, F. D. *Energy Fuels* **2014**, *28* (3), 2019–2026.

(54) Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill, H. H. *J. Mass Spectrom.* **2008**, *43* (1), 1–22.

(55) Gross, J. H. *Mass Spectrometry: A Textbook*, 2nd ed.; springer heidelberg dordrecht london new york, 2011.

(56) J. J. Thomson. *Philos. Mag.* **1897**, *44*, 293–316.

(57) Griffiths, J. *Anal. Chem.* **2008**, *80* (15), 5678–5683.

(58) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. *Rapid Commun. Mass Spectrom.* **1988**, *2* (8), 151–153.

(59) Karas, M.; Bachmann, D.; Hillenkamp, F. *Anal. Chem.* **1985**, *57* (14), 2935–2939.

(60) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49* (5), 2240–2249.

(61) Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984**, *88* (20), 4451–4459.

(62) Marshall, A. G.; Rodgers, R. P. *Proc. Natl. Acad. Sci.* **2008**, *105* (47), 18090–18095.

(63) Fakheri, R. J.; Javitt, N. B. *Steroids* **2012**, *77* (6), 575–577.

(64) Gaikwad, N. W. *Anal. Chem.* **2013**, *85* (10), 4951–4960.

(65) Griffiths, W. J.; Wang, Y. *Biochim. Biophys. Acta BBA - Mol. Cell Biol. Lipids* **2011**, *1811* (11), 784–799.

(66) Li, Y.; Liang, M.; Shu, J.; Zhang, P.; Sun, W. *Anal. Methods* **2014**, *6* (13), 4754–4760.

(67) Meljon, A.; Theofilopoulos, S.; Shackleton, C. H. L.; Watson, G. L.; Javitt, N. B.; Knölker, H.-J.; Saini, R.; Arenas, E.; Wang, Y.; Griffiths, W. J. *J. Lipid Res.* **2012**, *53* (11), 2469–2483.

(68) Wang, Y.; Karu, K.; Griffiths, W. J. *Biochimie* **2007**, *89* (2), 182–191.

(69) Wepf, A.; Glatter, T.; Schmidt, A.; Aebersold, R.; Gstaiger, M. *Nat. Methods* **2009**, *6* (3), 203–205.

(70) Solouki, T.; Paša-Tolić, L.; Jackson, G. S.; Guan, S.; Marshall, A. G. *Anal. Chem.* **1996**, *68* (21), 3718–3725.

(71) Mo, W.; Karger, B. L. *Curr. Opin. Chem. Biol.* **2002**, *6* (5), 666–675.

(72) Huang, Y.; S. Gelb, A.; D. Dodds, E. *Curr. Metabolomics* **2013**, *1* (4), 291–305.

(73) Munisamy, S. M.; Chambliss, C. K.; Becker, C. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (7), 1250–1259.

(74) Vismeh, R.; Humpula, J. F.; Chundawat, S. P. S.; Balan, V.; Dale, B. E.; Jones, A. D. *Carbohydr. Polym.* **2013**, *94* (2), 791–799.

(75) Izumi, Y.; Takimura, S.; Yamaguchi, S.; Iida, J.; Bamba, T.; Fukusaki, E. *J. Biosci. Bioeng.* **2012**, *113* (3), 412–419.

(76) Mandal, M. K.; Yoshimura, K.; Saha, S.; Yu, Z.; Takeda, S.; Hiraoka, K. *Anal. Chem.* **2014**, *86* (2), 987-992

(77) Sekar, R.; Kailasa, S. K.; Chen, Y.-C.; Wu, H.-F. *Chin. Chem. Lett.* **2014**, *25* (1), 39–45.

(78) Kuang, X.; Ma, Y.; Su, H.; Zhang, J.; Dong, Y.-B.; Tang, B. *Anal. Chem.* **2014**, *86* (2), 1277–1281.

(79) Kailemia, M. J.; Ruhaak, L. R.; Lebrilla, C. B.; Amster, I. J. *Anal. Chem.* **2014**, *86* (1), 196–212.

(80) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246* (4926), 64–71.

(81) Daub, C. D.; Cann, N. M. *Anal. Chem.* **2011**, *83* (22), 8372–8376.

(82) Wilm, M. *Mol. Cell. Proteomics MCP* **2011**, *10* (7), M111.009407.

(83) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2013**, *85* (1), 2–9.

(84) Makarov, A.; Denisov, E.; Kholomeev, A.; Balschun, W.; Lange, O.; Strupat, K.; Horning, S. *Anal. Chem.* **2006**, *78* (7), 2113–2120.

(85) March, R. E.; Todd, J. F. J. *Int. J. Mass Spectrom.* **2015**, *377*, 316–328.

(86) Subramaniam, R.; Östin, A.; Nygren, Y.; Juhlin, L.; Nilsson, C.; Åstot, C. *J. Mass Spectrom. JMS* **2011**, *46* (9), 917–924.

(87) Attygalle, A.; Ruzicka, J.; Varughese, D.; Sayed, J. *Tetrahedron Lett.* **2006**, *47* (27), 4601–4603.

(88) Zekavat, B.; Miladi, M.; Al-Fdeilat, A. H.; Somogyi, A.; Solouki, T. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (2), 226–236.

(89) Cuyckens, F.; Wassvik, C.; Mortishire-Smith, R. J.; Tresadern, G.; Campuzano, I.; Claereboudt, J. *Rapid Commun. Mass Spectrom.* **2011**, *25* (23), 3497–3503.

(90) Yates, J. R.; Eng, J. K.; Clauser, K. R.; Burlingame, A. L. *J. Am. Soc. Mass Spectrom.* **1996**, *7* (11), 1089–1098.

(91) Lee, H. H.; Hong, A.; Cho, Y.; Kim, S.; Kim, W. J.; Kim, H. I. *J. Am. Soc. Mass Spectrom.* **2015**, 1–10.

(92) Xian, F.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2012**, *84* (2), 708–719.

(93) Amorim Madeira, P. J.; A., P.; M., C. In *Fourier Transform - Materials Analysis*; Salih, S., Ed.; InTech, 2012.

(94) Brenton, A. G.; Godfrey, A. R. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (11), 1821–1835.

(95) Olsen, J. V.; de Godoy, L. M. F.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. *Mol. Cell. Proteomics MCP* **2005**, *4* (12), 2010–2021.

(96) McKenna, A. M.; Williams, J. T.; Putman, J. C.; Aeppli, C.; Reddy, C. M.; Valentine, D. L.; Lemkau, K. L.; Kellermann, M. Y.; Savory, J. J.; Kaiser, N. K.; Marshall, A. G.; Rodgers, R. P. *Energy Fuels* **2014**, *28* (4), 2454–2464.

(97) Bae, E.; Yeo, I. J.; Jeong, B.; Shin, Y.; Shin, K.-H.; Kim, S. *Anal. Chem.* **2011**, *83* (11), 4193–4199.

(98) Dirk Willem van Krevelen. *Fuel* **1950**, 269–284.

(99) McDaniel, E. W.; Martin, D. W.; Barnes, W. S. *Rev. Sci. Instrum.* **1962**, *33* (1), 2–7.

(100) Hill, H. H.; Siems, W. F.; St. Louis, R. H. *Anal. Chem.* **1990**, *62* (23), 1201A – 1209A.

(101) Ahmed, A.; Cho, Y. J.; No, M.; Koh, J.; Tomczyk, N.; Giles, K.; Yoo, J. S.; Kim, S. *Anal. Chem.* **2011**, *83* (1), 77–83.

(102) Fasciotti, M.; Lalli, P. M.; Klitzke, C. F.; Corilo, Y. E.; Pudenzi, M. A.; Pereira, R. C. L.; Bastos, W.; Daroda, R. J.; Eberlin, M. N. *Energy Fuels* **2013**, *27* (12), 7277–7286.

(103) Santos, J. M.; Galaverna, R. de S.; Pudenzi, M. A.; Schmidt, E. M.; Sanders, N. L.; Kurulugama, R. T.; Mordehai, A.; Stafford, G. C.; Wisniewski, A.; Eberlin, M. N. *Anal. Methods* **2015**, *7* (11), 4450–4463.

(104) Becker; Fernandez-Lima, F. A.; Russell, D. H. *Spectroscopy* **2009**, *24*, 38–42.

(105) Dwivedi, P.; Schultz, A. J.; Hill, H. H. *Int. J. Mass Spectrom.* **2010**, *298* (1-3), 78–90.

(106) Lutz, N. W.; Sweedler, J. V.; Wevers, R. A. *Methodologies for Metabolomics: Experimental Strategies and Techniques*; Cambridge University Press, 2013.

(107) Zinnel, N. F.; Pai, P.-J.; Russell, D. H. *Anal. Chem.* **2012**, *84* (7), 3390–3397.

(108) McLean, J. A.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H. *Int. J. Mass Spectrom.* **2005**, *240* (3), 301–315.

(109) Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. *Rapid Commun. Mass Spectrom.* **2004**, *18* (20), 2401–2414.

(110) Giles, K.; Williams, J. P.; Campuzano, I. *Rapid Commun. Mass Spectrom.* **2011**, *25* (11), 1559–1566.

(111) Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. *Nat. Chem.* **2014**, *6* (4), 281–294.

(112) Munisamy, S. M.; Chambliss, C. K.; Becker, C. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (7), 1250–1259.

(113) Lavanant, H.; Tognetti, V.; Afonso, C. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 572–580.

(114) Harvey, D. J.; Scarff, C. A.; Edgeworth, M.; Crispin, M.; Scanlan, C. N.; Sobott, F.; Allman, S.; Baruah, K.; Pritchard, L.; Scrivens, J. H. *Electrophoresis* **2013**, *34* (16), 2368–2378.

(115) Hariharan, C. B.; Baumbach, J. I.; Vautz, W. *Anal. Chem.* **2010**, *82* (1), 427–431.

(116) Maurer, M. M.; Donohoe, G. C.; Valentine, S. J. *Analyst* **2015,** *140*, 6782–6798.

(117) Vamvuka, D. *Int. J. Energy Res.* **2011**, *35* (10), 835–862.

(118) Yaman, S. *Energy Convers. Manag.* **2004**, *45* (5), 651–671.

(119) Mohan, D.; Pittman, Charles U.; Steele, P. H. *Energy Fuels* **2006**, *20* (3), 848–889.

(120) Rogers, J. G.; Brammer, J. G. *Biomass Bioenergy* **2012**, *36*, 208–217.

(121) Mettler, M. S.; Vlachos, D. G.; Dauenhauer, P. J. *Energy Environ. Sci.* **2012**, *5* (7), 7797–7809.

(122) Kanaujia, P. K.; Sharma, Y. K.; Agrawal, U. C.; Garg, M. O. *TrAC Trends Anal. Chem.* **2013**, *42*, 125–136.

(123) Mullen, C. A.; Strahan, G. D.; Boateng, A. A. *Energy Fuels* **2009**, *23* (5), 2707–2718.

(124) Luik, H.; Luik, L.; Tiikma, L.; Vink, N. *J. Anal. Appl. Pyrolysis* **2007**, *79* (1–2), 205–209.

(125) Chen, T.; Deng, C.; Liu, R. *Energy Fuels* **2010**, *24* (12), 6616–6623.

(126) Cole, D. P.; Smith, E. A.; Dalluge, D.; Wilson, D. M.; Heaton, E. A.; Brown, R. C.; Lee, Y. J. *Fuel* **2013**, *111*, 718–726.

(127) Sudasinghe, N.; Dungan, B.; Lammers, P.; Albrecht, K.; Elliott, D.; Hallen, R.; Schaub, T. *Fuel* **2014**, *119*, 47–56.

(128) Olcese, R.; Carré, V.; Aubriet, F.; Dufour, A. *Energy Fuels* **2013**, *27* (4), 2135–2145.

(129) Tessarolo, N. S.; Silva, R. C.; Vanini, G.; Pinho, A.; Romão, W.; de Castro, E. V. R.; Azevedo, D. A. *Microchem. J.* **2014**, *117*, 68–76.

(130) Venkatakrishnan, V. K.; Degenstein, J. C.; Smeltz, A. D.; Delgass, W. N.; Agrawal, R.; Ribeiro, F. H. *Green Chem.* **2014**, *16* (2), 792–802.

(131) Chiaberge, S.; Leonardis, I.; Fiorani, T.; Cesti, P.; Reale, S.; DeAngelis, F. *Energy Fuels* **2014**, *28*, 2019–2026.

(132) Zhu, X.; Chen, Y.; Subramanian, R. *Anal. Chem.* **2014**, *86* (2), 1202-1209.

(133) Zekavat, B.; Szulejko, J. E.; LaBrecque, D.; Olaitan, A. D.; Solouki, T. *Rapid Commun. Mass Spectrom.* **2014**, *28* (2), 230–238.

(134) Hariharan, C.; Baumbach, J. I.; Vautz, W. *Int. J. Ion Mobil. Spectrom.* **2009**, *12* (2), 59–63.

(135) Becker, C.; Fernandez-Lima, F. A.; Gillig, K. J.; Russell, W. K.; Cologna, S. M.; Russell, D. H. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (6), 907–914.

(136) Dhungana, B.; Becker, C.; Zekavat, B.; Solouki, T.; Hockaday, W. C.; Chambliss, C. K. *Energy Fuels* **2015**, *29* (2), 744–753.

(137) Verevkin, S. P.; Zaitsau, D. H.; Emel'yanenko, V. N.; Yermalayeu, A. V.; Schick, C.; Liu, H.; Maginn, E. J.; Bulut, S.; Krossing, I.; Kalb, R. *J. Phys. Chem. B* **2013**, *117* (21), 6473–6486.

(138) Gawdzik, B.; Osypiuk, J. *J. Chromatogr. A* **2000**, *898* (1), 13–21.

(139) Demoranville, L. T.; Houssiau, L.; Gillen, G. *Anal. Chem.* **2013**, *85* (5), 2652–2658.

(140) Zamora, D.; Alcalà, M.; Blanco, M. *Anal. Chim. Acta* **2011**, *708* (1–2), 69–74.

(141) Katzenmeyer, B. C.; Cool, L. R.; Williams, J. P.; Craven, K.; Brown, J. M.; Wesdemiotis, C. *Int. J. Mass Spectrom.* **2015**, *378*, 303–311.

(142) Campbell, J. L.; Zhu, M.; Hopkins, W. S. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (9), 1583–1591.

(143) Schneider, B. B.; Covey, T. R.; Nazarov, E. G. *Int. J. Ion Mobil. Spectrom.* **2013**, *16* (3), 207–216.

(144) Gidden, J.; Wyttenbach, T.; Jackson, A. T.; Scrivens, J. H.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122* (19), 4692–4699.

(145) Knapman, T. W.; Berryman, J. T.; Campuzano, I.; Harris, S. A.; Ashcroft, A. E. *Int. J. Mass Spectrom.* **2010**, *298* (1–3), 17–23.

(146) Giles, K.; Wildgoose, J. L.; Langridge, D. J.; Campuzano, I. *Int. J. Mass Spectrom.* **2010**, *298* (1–3), 10–16.

(147) Nikolaev, E. N.; Kostyukevich, Y. I.; Vladimirov, G. N. *Mass Spectrom. Rev.* **2014**, doi:10.1002/mas.21422.

(148) Staš, M.; Chudoba, J.; Kubička, D.; Pospíšil, M. *Energy Fuels* **2015**, *29* (5), 3233–3240.

(149) Sharma, L. N.; Chambliss, C. K. Identification and quantitation of potential fermentation inhibitors in biomass pretreatment hydrolysates using highperformance liquid chromatography in combination with ultraviolet detection and tandem mass spectrometry, Baylor University: Waco, Tex, 2009.