



**Michigan
Technological
University**

Michigan Technological University
Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's Reports

2019

METHOD CONSIDERATIONS FOR COMPOUND IDENTIFICATION IN COMPLEX MIXTURES USING ELECTROSPRAY IONIZATION ULTRAHIGH RESOLUTION MASS SPECTROMETRY


Tyler Leverton
Michigan Technological University, talevert@mtu.edu

Copyright 2019 Tyler Leverton

Recommended Citation

Leverton, Tyler, "METHOD CONSIDERATIONS FOR COMPOUND IDENTIFICATION IN COMPLEX MIXTURES USING ELECTROSPRAY IONIZATION ULTRAHIGH RESOLUTION MASS SPECTROMETRY", Open Access Master's Thesis, Michigan Technological University, 2019.
<https://doi.org/10.37099/mtu.dc.etr/811>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etr>

 Part of the [Analytical Chemistry Commons](#), [Environmental Chemistry Commons](#), and the [Organic Chemistry Commons](#)

METHOD CONSIDERATIONS FOR COMPOUND IDENTIFICATION IN COMPLEX
MIXTURES USING ELECTROSPRAY IONIZATION ULTRAHIGH RESOLUTION
MASS SPECTROMETRY

By

Tyler A. Leverton

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Chemistry

MICHIGAN TECHNOLOGICAL UNIVERSITY

2019

© 2019 Tyler Leverton

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Chemistry.

Department of Chemistry

Thesis Advisor: *Dr. Lynn R. Mazzoleni*

Committee Member: *Dr. Marina Tanasova*

Committee Member: *Dr. Daisuke Minakata*

Department Chair: *Dr. John A. Jaszczak.*

Table of Contents

Acknowledgements.....	v
Abstract.....	vi
1 Introduction.....	1
1.1 Atmospheric Aerosol.....	1
1.2 Ultrahigh Resolution Mass Spectrometry	2
1.3 Electrospray Ionization.....	4
1.4 Data Analysis of Complex Mixtures – Mass Spectrometry Informatics.....	7
1.5 Interpretation of Electrospray Ionization-Ultrahigh Resolution Mass Spectra and Sample Preparation Considerations	11
1.5.1 Extent of Hydrogen Deuterium Exchange as a Potential Indicator of Functional Group Composition	13
1.5.2 Solid Phase Extraction and Ammonium Artifacts	15
1.5.3 Reactive Artifacts with the Use of Methanol as a Solvent	17
1.6 Thesis Structure.....	18
2 Materials and Methods.....	20
2.1 Sample Preparation.....	20
2.1.1 Hydrogen-Deuterium Exchange Sample Preparation.....	20
2.1.1.1 Acid Mix Sample Preparation.....	20
2.1.1.2 Mesquite Liquid Smoke Sample Preparation	21
2.1.2 NH ₄ OH Sample Preparation	21
2.1.3 Methanol Artifact Sample Preparation	23
2.2 Ultrahigh Resolution Orbitrap Elite Mass Spectrometry Analysis	23
2.3 Data Processing and Molecular Formula Assignment for Liquid Smoke Samples.....	24
2.3.1 Molecular Formula Assignment Parameters.....	26
2.3.1.1 Hydrogen-Deuterium Exchange	26
2.3.1.2 NH ₄ OH Artifact	27
2.3.1.3 Methanol Artifact.....	28
2.3.2 Reducing Ambiguity and Data Size.....	28
3 Characterization of Labile Hydrogen in Mesquite Liquid Smoke Using Hydrogen- Deuterium Exchange.....	30
3.1 Hydrogen-Deuterium Exchange of Acid Mix	30
3.2 Negative ESI Analysis of Mesquite Liquid Smoke.....	35
3.2.1 Preliminary Molecular Formula Composition	35

3.2.2	Filtering Formulas.....	38
3.2.3	Trends in HDX-Capable Species.....	45
3.3	Positive ESI Analysis of Mesquite Liquid Smoke	50
3.3.1	Preliminary Molecular Formula Composition.....	50
3.3.2	Filtering Formulas.....	54
3.3.2.1	Trends in HDX-Capable Species.....	61
3.4	Comparison of Negative and Positive HDX-Capable Species.....	71
4	Characterization of Ammonium Artifacts Formed Using a 2-Step Solid-Phase Extraction Method with Liquid Smoke.....	77
4.1	Negative ESI Analysis.....	77
4.1.1	Initial Molecular Formula Assignment.....	77
4.1.2	Ambiguity and Quality Assurance.....	81
4.1.3	Composition of ¹⁵ N Molecular Formulas	83
4.1.4	Potential Adduct and Reactive Artifact Pathways.....	89
4.2	Positive ESI Analysis	100
4.2.1	Initial Molecular Formula Assignment.....	100
4.2.2	Ambiguity and Quality Assurance.....	101
4.2.3	Composition of ¹⁵ N Molecular Formulas	105
4.2.4	Potential Adduct and Reaction Artifact Pathways.....	110
5	Characterization of Artifacts Formed in Methanol-Solvated Liquid Smoke Using Electrospray Ionization	117
5.1	Preliminary Molecular Formula Composition	117
5.2	Preliminary Filtering of Formulas.....	122
5.3	Potential Pathways of Artifact Formation.....	129
6	Conclusions.....	140
6.1	Hydrogen-Deuterium Exchange Conclusions.....	140
6.2	NH ₄ OH Artifact Conclusions.....	142
6.3	Methanol Artifact Conclusions	143
7	Reference List.....	145

Acknowledgements

I would like to thank Dr. Lynn Mazzoleni for her extensive guidance and teachings that challenged me and allowed me to succeed. I would also like to thank Simeon Schum, Matt Brege, and Maryam Khaksari for their assistance and advice with instrumentation and data analysis. I would like to give special thanks to Corrin Davis for her love and support, as without her, this work would not have been possible.

In addition, I would like to thank those in the Department of Chemistry who gave me the knowledge and opportunities for growth and for financial assistance. Thank you to Drs. Marina Tanasova and Daisuke Minakata as well for taking the time to offer invaluable feedback on this work.

Abstract

Biomass burning aerosols are highly complex organic mixtures of thousands of components with consequences for global climate. Complex mixture component identification requires accurate mass measurement capability to separate components on a milli-Dalton scale, frequently using ultrahigh resolution mass spectrometry with electrospray ionization. Certain sample preparations and the ionization process may introduce artifacts that obscure the composition of the sample. Two method considerations were explored using isotopically labeled $^{15}\text{NH}_4\text{OH}$ and MeOH-d_3 to track artifact formation in biomass burning samples. Informatics techniques and a custom molecular formula assignment software were used to identify the isotopic atoms in artifact products. Sample preparation with NH_4OH was found to significantly alter the detected complex mixture composition, potentially by NH_4^+ adduction and by reactions between NH_3 and carbonyls. Solvation in MeOH likely induced artifact formation by converting carbonyls to esters and acetals/hemiacetals for non-aromatic species. In addition, Hydrogen-Deuterium exchange using MeOH-d_1 was studied to infer the presence of amino and hydroxyl groups, and estimate the number of carboxyl functional groups. The results have important implications for ultrahigh resolution mass spectrometry analyses of complex environmental samples and their labile H content.

1 Introduction

1.1 Atmospheric Aerosol

Atmospheric aerosols are widely accepted as significant contributors to the earth's changing climate (IPCC 2014, Shiraiwa et al. 2017). These small particles are typically aggregates of organic or inorganic molecules that enter the atmosphere through various means, such as industrial coal burning, sand and dust erosion, and biomass burning (*e.g.* forest fires) (Remer and Kaufman 1998; IPCC 2014; Cohen et al. 2017). Biomass burning is considered to be one of the largest contributors of particulate carbon in the atmosphere and is becoming more severe due to increases in the number and severity of burning events in the last few decades (Westerling et al. 2006). Thus, it is very important to understand how biomass burning organic aerosol specifically contributes to the global radiative budget.

Understanding the effect of aerosol on climate requires thorough analysis of its composition. However, aerosol samples are highly complex mixtures potentially consisting of thousands of unique constituents that make their analyses challenging (Mazzoleni et al. 2010; Mazzoleni et al 2012). Frequently the analyses tend to be sample limited due to the relatively small volumes of aerosol collected during campaigns due to a competing interest for diurnal study. Therefore, an abundant, affordable surrogate material is desirable to optimize method development with respect to accurate species detection before applying the method to an obtained sample. Although there is no universal standard for conducting aerosol analysis, there have been attempts to use similarly complex mixtures for method development. One such mixture is Suwanee River

Fulvic Acid (SRFA), an aqueous organic matter sample originally considered important due to early studies indicating its similarity to aerosol organic matter (Decesari et al. 2003; Perdue 2013). This similarity is likely due to the influence of plant matter (e.g. lignin decomposition products), but there are significant differences due to terpene oxidation products in aerosol organic matter (Mazzoleni et al. 2012).

A suitable alternative to SRFA to represent biomass burning aerosol is liquid smoke and can be used for method development experiments. Liquid smoke is a food additive prepared using a low temperature burning method to produce smoke from various wood types (mesquite, hickory). The smoke is captured and concentrated in water (B&G Foods, Inc. 2019; Montazeri et al. 2012). Since liquid smoke is a product of wood burning, it is expected to contain biomass combustion species such as lignin and cellulose degradation products like phenols. Thus, liquid smoke is a suitable surrogate for biomass burning samples, and is the sample used herein for method development and evaluation of the potential for method artifacts. In addition, its aqueous nature makes it easier to work with than ambient aerosol samples.

1.2 Ultrahigh Resolution Mass Spectrometry

To understand the detailed compositions of complex organic mixtures, it is necessary to use instrumentation that can separate components as much as possible. Complex mixtures, including liquid smoke and atmospheric aerosol, can contain thousands of unique molecular formulas (Brege et al. 2018). For this reason, ultrahigh resolution mass spectrometry (UHRMS) is commonly used. Ultrahigh resolution mass

spectrometers are instruments that separate ions based on their mass to charge ratio (m/z) for exact measurement. In contrast, low resolution mass spectrometers measure components with accuracy of approximately one nominal mass unit and ultrahigh resolution mass spectrometers separate masses with high decimal (mDa) precision. This makes it possible to differentiate ions within the same nominal mass. For example, the ions $C_3H_5O^+$ and $C_4H_9^+$ are both nominally m/z 57. However, using a high-resolution mass spectrometer, these two ions would yield two separate signals at m/z 57.07043 and m/z 57.03404 m/z . The two species have a mass difference of 0.03639 Da, or 90 ppm (Hoffmann 2005). Exact mass measurements with less than 3 ppm error can be used to determine the molecular formula (MF) of an unknown ion.

Mass resolving power defines the separation capability of a mass analyzer. The ultrahigh resolution Orbitrap Elite MS is capable of analyzing ions at a resolving power of 240,000 at m/z 400. There are two electrodes in the Orbitrap Elite mass analyzer. The inner electrode is spindle-shaped and the outer electrode is bell-shaped. The shape of the electrodes provides a higher electrical potential at the poles and a lower electrical potential in the center making it highly effective at separating minute differences in m/z . This potential separates packets of ions by the amount of time required for the ions to move back and forth along the length of the electrode (Michalski et al. 2012). The ion signal is recorded as an image current in the time domain, and then a Fourier transformation is applied to convert the time domain signal to an m/z value where mDa differences (< 1 ppm) can be resolved. For example, the difference between C_3 and SH_4 (3.4 mDa, or 0.85 ppm at m/z 400) can be detected with the Orbitrap Elite. However,

ultrahigh resolution cannot separate structural isomers, as they have the same exact mass. Structural characterization is a continuing challenge in mass spectrometry (Ramanathan et al. 1998; Yasmeen et al. 2011; Isaacman et al. 2012).

1.3 Electrospray Ionization

Ionization is one of the most important aspects to analyzing compounds with mass spectrometry and methods can be hard or soft. Electron ionization, for example, is often too energetic to preserve the analyte structures, and therefore is not suitable for analyzing complex mixtures, as the fragmentation would further complicate the mass spectra. Thus, softer ionization methods with little to no fragmentation are preferred. One common ionization process is known as electrospray ionization (ESI), which is known for its versatility in analyzing a wide variety of analytes. (Fenn et al. 1989; Smith et al. 1990; Kujawinski et al. 2002). ESI is an electrochemical process that produces gas-phase ions by desolvation. Samples are ionized in solution using a strong 2-5 kV power supply connected to the spray needle, where oxidation or reduction occurs. At the opposite end of the atmospheric pressure compartment, reduction or oxidation takes place at the oppositely charged plate, which has an aperture that guides ions into the mass spectrometer. A potential difference acts to pull solvated positive or negative charges away from the tip of the spray needle. This creates a cone-like shape at the tip known as a Taylor cone. When the solution at the end of the Taylor cone overcomes its surface tension a plume of solvent droplets with excess positive or negative charge is created. These charges reside on the surface of the droplets to minimize electrostatic repulsion. As

the solvent evaporates from the droplet the excess positive charges are brought closer together until the Rayleigh limit is reached (Taflin et al. 1989). The charge repulsion then overcomes the surface tension and forms a series of smaller droplets. This process continues until the solvent has evaporated, leaving behind individual gas-phase ions. These ions (*e. g.* $[M + H]^+$ or $[M - H]^-$) then enter the mass spectrometer for analysis.

A stable electrospray plume is necessary in order to maximize the signal obtained from the ESI process. This can be optimized by modifying several parameters which can be adjusted based on the sample. The ability for analytes to desolvate is perhaps most important and can be modified with several different parameters, such as a drying sheath gas (N_2), flow rate, and heat. The sample solvents also play a large role. Most ESI analyses utilize some combination of organic solvent and water, although the percentages can vary based on the sample composition and other parameters (Cech and Enke 2001a; Henriksen et al. 2005; Flerus et al. 2011; Novotny et al 2014). High percentages of organic solvent are desirable to increase the rate of desolvation. Typical solvents for electrospray are methanol (MeOH) and acetonitrile (ACN), which are used due to their small size, volatility, and moderate polarity. This makes them suitable for solvating many varieties of analytes while being easily evaporated (Liigand et al. 2014). Some percentage of water is also necessary for a stable electrospray, as its high surface tension facilitates the formation of a Taylor cone. It also contributes to the reduction-oxidation process for excess charge. However, too much water can become too difficult to desolvate and greatly reduce the signal (Henriksen et al. 2005). Thus, there is a fine balance that must exist between organic and aqueous solvent percentages.

Small, weak acids or their conjugate bases can also be added to the solvent to enhance ionization. These species can ionize efficiently and enhance the conductivity of the solution, improving the ionization of analytes and the overall electrospray stability (Cech and Enke 2001b). Acetic and formic acid and their conjugates have been frequently used for this purpose. The acid or salt used depends on the sample composition and the electrospray polarity.

The two electrospray polarities, positive and negative mode, ionize species differently and thus can be used to obtain knowledge on different species in a sample (Lin et al. 2018). In the negative mode, the analyte species are ionized via deprotonation and thus analytes must have labile protons that can be abstracted from the analyte. These are typically acidic species such as carboxylic acids, although it has been shown that the ESI response is not entirely dependent on pK_a , but on analyte hydrophobicity as well due to easier desolvation (Henriksen et al. 2005). Thus, the negative ion mode is used primarily to obtain information on acidic species containing C, H, and O atoms (CHO species), as well as on oxidized N or S species. Conversely, positive electrospray ionizes species by cation adduction. Cations can include H^+ , alkali metal ions such as Na^+ or K^+ , or polyatomic cations such as NH_4^+ , making positive mode analyses more difficult. Na^+ adducts are very common in positive ESI, and are problematic because they further complicate the analysis of complex mixtures. Na^+ adducts have been shown to be present more frequently in O-containing compounds than in N-containing compounds (Krueve et al. 2013). NH_4^+ adducts are understood less than others due to a limited number of detailed studies although these adducts are highly relevant to atmospheric analyses

because of the prevalence of sulfate (Galloway et al. 2009; Lin et al. 2015; Hawkins et al. 2018). However, results herein offer some evidence that NH_4^+ adducts and reaction artifacts may also occur in some sample preparation conditions (See Chapter 3). Regardless of the adduct formation considerations, positive ESI is a reliable ionization method for basic analytes such as amines or other reduced N-containing species (CHNO) (Reemtsma et al. 2006; Ehrmann et al. 2008).

Analyte deprotonation and adduct formation can be used to ionize a wide range of molecules. However, unfunctionalized molecules such as polycyclic aromatic hydrocarbons or alkanes are not suitable, as they are highly resistant to being ionized since ESI produces even electron species. Other ionization methods, such as atmospheric pressure photoionization (APPI), may be better suited for these analytes due to different ionization pathways (Purcell et al. 2007). The studies contained in this thesis, however, focus solely on species detected using ESI with the knowledge that other species may be present but are undetected.

1.4 Data Analysis of Complex Mixtures – Mass Spectrometry Informatics

Once ions have been created and mass spectrometric information has been collected and exported as a mass list, the data needs extensive post processing to determine the identities of the masses. It is desired to assign MF to each ion based on their exact mass values to understand the composition of the sample, such as fatty acid recalibrants (Sleighter et al). However, it is impractical to manually assign MF to each detected ion within a complex mixture due to the large number of ions. As a result,

computational and data science techniques are utilized to interpret ultrahigh resolution mass spectra of complex mixtures, assign MF to those ions, and understand the nature and characteristics of those MF and subsequently the complex sample.

In this data-science driven approach, several steps are required to ensure the detected ions are assigned accurate MF for further analysis. Although the details in each step vary among programs and methods, the core of each method is relatively similar. First, it is necessary to recalibrate the mass list (Sleighter et al. 2008) by selecting preliminary MF within the sample. Mass measurement error is inevitable in any electronic instrument due to noise and electronic drift (Wu and McAllister 2003; Makarov et al. 2006; Brenton and Godfrey 2010). For example, in ultrahigh resolution mass spectrometry, higher m/z ions tend to have higher error due to the reduced mass resolution and tend to have a larger number of possible formulas (Ohno and Ohno 2013). Mass recalibration is done to minimize the overall mass error. This is done by using measured masses present in the sample. If the sample does not have a known set of masses, recalibration is performed by selecting a set of the most reliable ions in the mass spectrum, typically common CHO MF for biomass burning-like complex mixtures, or known hydrocarbons in petroleum-based samples (Sleighter et al. 2008; Kozhinov et al. 2013). The recalibration process greatly increases the reliability of MF assignments.

Molecular formula assignment is then performed on the recalibrated masses to assign the most reliable formula or formulas. This process becomes increasingly complicated at higher m/z values, where there can be several theoretically plausible MF for one given mass. Allowing multiple heteroatoms increases the number of possible MF,

as does more isotopes, though has not been thoroughly explored (Koch et al. 2007; Ohno and Ohno 2013). When multiple plausible MF are possible they are defined as “ambiguous formulas,” (Schum et al. 2019) and present a key challenge to the data interpretation (See Section 2.3.2). Although approaches to reducing MF ambiguity differ between methods, programs, and samples, a common goal is to remove chemically unreasonable formulas. The plausibility of formulas is based on various measures, such as atomic valence, double bond equivalents (DBE), and the number of heteroatoms (Kind and Fiehn 2007; Ohno and Ohno 2013; Schum 2019). In addition, complex mixtures resulting from one common source are generally composed of groups of molecular formulas that are linked through base molecular units. For example, in biomass burning samples, it is expected that there will be a wide range of compounds with the same number of O, but that differ by some number of methylene (-CH₂-) groups (Kendrick 1963; Stenson et al. 2003; Smith et al. 2009). This is an additional quality assurance (QA) step than can be used to simplify the data and ensure MF accuracy (Schum 2019).

MF assignment software, such as Composer64 (Sierra Analytics), have various methods for assigning MF to ultrahigh resolution mass spectra and reducing ambiguity (Stanz 2015, Gavard et al. 2017). Composer64 reduces ambiguity by selecting the most hydrocarbon-like MF, although the parameters for MF assignment are relatively flexible and customizable. The software is reliable for most applications, but is limited with respect to isotope assignments and transparent ambiguity decisions. Methods for performing MF assignment, choosing the most likely MF, and removing unlikely assignments are also poorly understood or not public knowledge. This can make some

assignments difficult to justify or improve. In addition, these programs typically require specific licensing agreements.

Due to the aforementioned reasons, a new method for molecular formula assignment was developed by Schum (2019). The newly designed molecular assignment software, MFAssignR, is available via GitHub (Schum et al. 2019). MFAssignR was written in R with customizable and transparent methods for MF assignment with full ambiguity. The functions perform a semi-automatic recalibration and assign MF to ions with the ability to report all chemically reasonable ambiguous MF assignments. The latter capability makes it suitable for isotope assignments, which is important in the scope of this work; MFAssignR is capable of assigning MF with ^{13}C , ^{34}S , ^{15}N , and D. Due to this versatility, MFAssignR was the method of molecular formula assignment used in this work and the primary method of recalibration alongside Composer64.

Assigned MF can be interpreted using functions within RStudio as well through data manipulation and extraction functions (Wickham 2016). Since data sets from the ultrahigh resolution MS of complex mixtures and their assigned MF are quite large (upwards of 5000 assigned MF with nearly 50 descriptor variables), extensive data science is required to provide appropriate molecular insight, which can be defined as ‘Mass Spectrometry Informatics’.

1.5 Interpretation of Electrospray Ionization-Ultrahigh Resolution Mass Spectra and Sample Preparation Considerations

With the highly complicated composition of organic mixtures, it is difficult to ensure analytes are properly characterized, even when using ultrahigh resolution mass spectrometry to separate compounds nearly identical in mass. In order to properly analyze and identify individual components in complex organic mixtures, it is exceedingly important that samples are prepared in ways that reliably extract these components with as little interference as possible. There are many important factors that can enhance analytical accuracy and reliability, including solvent factors (e.g. amount of organic solvent, acidity), instrumental parameters (e.g. electrospray temperature, voltage), and sample preparation techniques (e.g. collection and extraction methods). These factors are not universal, and vary based on the sample type and source.

ESI in particular is of great interest due to the nature of the ionization process. Knowledge of how the ESI-MS sample analysis process affects analyte ions is somewhat limited due to an incomplete understanding of how the excess charge in the solution may introduce harsh conditions for the sample ions (Zhou et al. 2002). It is theorized that the pH of the solution changes drastically during the electrospray process and these effects have been linked to accretion interactions among protein molecules and solvents (Hossain and Konermann 2006). However, less is known about how ESI may force interactions between solvent molecules and smaller organic analytes in complex mixtures to create artifacts or if artifacts arise from sample preparation methods. In either case,

due to the extensive complexity of these mixtures, it is difficult to interpret the mass spectra and determine the effects without the MS informatics tools previously mentioned.

Since there are several methods that are widely utilized for the ESI-MS analysis of aerosols and complex mixtures such as solid-phase extraction and solvation in methanol, it is important to know how or if these methods impact the sample composition. Knowledge of these interactions, such as the exchange rate of labile protons in sample components, can also be used alongside data science methods to better understand the sample components. In addition, it is beneficial to know how or if the electrospray process drives any of these interactions. Ultimately, the purpose of this work is to offer considerations into how different methods of sample preparation may impact the analyses of complex organic mixtures and identification of their components. The steps to doing so are threefold: 1) Understand the extent to which hydrogen-deuterium exchange (HDX) occurs in MeOH- d_1 solvent, and use it to characterize labile H in liquid smoke species; 2) Investigate the extent to which MeOH artifacts are formed in liquid smoke during either the sample preparation or electrospray processes when added as the solvent; and 3) Investigate the extent of NH_4OH artifact or adduct formation during solid-phase extraction (SPE) of liquid smoke. The three aforementioned goals are described in detail in the following sections.

1.5.1 Extent of Hydrogen Deuterium Exchange as a Potential Indicator of Functional Group Composition

Given the variable nature of complex organic mixtures, it can be challenging to confirm the presence of particular functional groups or structural moieties without performing additional experiments. Assigning a molecular formula reveals little to no information about the types of bonds in the molecule, and may be composed of different isomers (Ramanathan et al. 1998; Yasmeen et al. 2011; Isaacman et al. 2012; Zark et al. 2017). However, functional groups with labile hydrogen atoms (*i.e.* hydroxyls, carboxyls, and amines) may be discerned via hydrogen exchange. In the presence of a labile deuterated solvent, such as D₂O or MeOH-d₁, acidic hydrogens within a sample can be replaced by deuterium, known as hydrogen-deuterium exchange (HDX).

This phenomenon occurs readily with non-deuterated solvents, but is not detectable since there is no change in the molecular characteristics. With deuterium, the occurrence of exchange can be detected by measuring changes in mass after deuterium has been introduced. Specifically, if one HDX occurs, the new compound will have a mass increase of 1.006277 Da. Ultrahigh resolution mass spectrometry and data science techniques are useful tools for studying the extent of HDX for complex mixtures, because they provide exact mass measurements and interpretation of those measurements.

HDX has been used in mass spectrometric analyses to a somewhat limited extent. The majority of the literature on HDX is focused on protein analysis, because it can be used to study protein conformations (Yan and Maier 2009; Oganessian 2018; Kostyukevich et al. 2018). Kostyukevich *et al.* reported two studies on the use of HDX in

complex mixture analysis (Kostyukevich et al. 2013; Kostyukevich, Kononikhin, Popov, & Nikolaev, 2013), using Fourier-transform ion cyclotron resonance (FT-ICR) MS data as well as both deuterated solvents and a deuterated atmosphere. The benefit of using a deuterated atmosphere was that it ensured the maximum possible depth of HDX, that is, every labile H in a molecule could be exchanged. However, this technique requires knowledge of the mixture components of the mixture to generate molecular formulas. Conversely, the use of a deuterated solvent has a low depth and cannot relay information on the total number of labile hydrogens in a molecule due to the rapid back-exchange of deuterated molecules with water in the gas phase in ESI (Kostyukevich et al. 2013). However, the low depth virtually ensures the non-deuterated molecule will be detectable and can make it simpler to ensure correct formulas with an automatic formula assignment software without requiring knowledge of sample components.

To the extent of our knowledge, there is little in the literature regarding the use of HDX as a technique to summarize the potential functional group composition of a complex organic mixture. One study by Stenson et al. (2014) used HDX to differentiate isomers in different ion-molecule reaction experiments using SRFA. Kostyukevich *et al.* (2013) differentiated between exchangeable and non-exchangeable O-containing groups based on their deuterated atmosphere technique, but they did not distinguish between carboxyls and hydroxyls, and they did not study other molecular parameters such as double bond equivalents (DBE) that could relate to the number of carboxyl groups. In addition, there is little knowledge regarding the molecular characteristics of liquid smoke. The goal of this investigation was to investigate the probable extent to which hydrogen-

deuterium exchange (HDX) occurs when liquid smoke is solvated in hydroxyl-deuterated MeOH (MeOH-d₁) using ultrahigh resolution MS, and determine if it can be used to highlight certain compounds and/or functional groups present in a sample.

1.5.2 Solid Phase Extraction and Ammonium Artifacts

One of the ubiquitous concerns of preparing any complex environmental sample is ensuring that a representative set of species are collected and analyzed. Given the broad ranges of compounds, it is nearly impossible to attain a complete representation of the sample. For example, a majority of compounds in mesquite liquid smoke (MLS) are water soluble, but there are also some species that require less polar solvents and different ionization methods despite the versatility of ESI-MS. Even among the water soluble components, there is a wide range of sizes, polarities, and elemental compositions that can make analysis difficult.

The aerosol chemistry community has attempted to mitigate this difficulty via many different routes. One such method is to prepare and solvate samples using a solid phase extraction (SPE) technique (Lin et al 2012). This chromatography technique separates components using a compact SPE cartridge based on differences in interaction between the mobile and stationary phases. Similar to column chromatography, complex mixtures can be separated using different SPE elution conditions. After extracting components from a sample filter and washing away aqueous salts that can interfere with the electrospray process, a small amount of HCl is added to load the sample to neutralize organic anions and enhance their retention. Then organic material bound to the SPE stationary phase is washed off with an organic/aqueous solvent mix such as MeOH (L1)

then a small amount (0.3%) of NH_4OH in elution 2 (L2). The ammonium hydroxide is used to recover less soluble species that are more strongly bound to the stationary phase.

Although this procedure was proven to be effective for recovering a large amount of organic material (Lin et al. 2012), there is a concern that ammonium in the SPE mobile phase interacts with sample analytes. Two potential interactions that may occur. First, residual NH_4^+ may adduct to analytes during the ESI process, much like H^+ or Na^+ . This may yield incorrect molecular formula assignments. Second, NH_3 , which is in equilibrium with NH_4^+ , may react with carbonyl compounds through addition or elimination reactions to form amines, imines, or imidazoles. It has been shown previously that the presence of NH_3 in the atmosphere can form imine or imidazole-based artifacts within aerosol (Galloway et al 2009; De Haan et al. 2011; Teich et al. 2016; Hawkins et al. 2018) but this reactivity has not been well characterized for the SPE process. The reactive artifact is of particular concern because it both alters the molecular structure and composition of complex samples, as well as artificially introduces N.

If these reactive artifacts occur from the addition of NH_4OH to samples, the presence of acid could either accentuate or dampen the reaction process depending on how it interacts with species in solution. If the pH is low enough, the equilibrium shifts to favor the presence of NH_4^+ , making it inert as a nucleophile. However, if the acid interacts with the carbonyl, it could increase the electrophilicity of the carbonyl carbon and promote a reaction with the NH_3 (Reusch 1999). The reaction pathway depends on the type of carbonyl. NH_3 could react with ketones or aldehydes to form a hemiacetal then an acetal, or it could react with esters to form amides.

The purpose of this investigation was to characterize and quantify the extent of adduct or reaction artifact formation from the second elution mobile phase used in the SPE process with mesquite and hickory liquid smoke as test samples. This was done by analyzing both liquid smoke samples with regular and ^{15}N -labeled ammonium hydroxide. This study has important implications for the accurate identification of components in complex organic mixtures.

1.5.3 Reactive Artifacts with the Use of Methanol as a Solvent

The solvent choice for ESI analysis is exceedingly important for accurate identification of complex sample components. Since analytes must be solvated for the ESI process, the solvent must have a balance of solvating a vast range of analytes and being sufficiently volatile to be removed. One such solvent is MeOH, which has been shown in several experiments to result in a more stable electrospray and thus an overall higher response compared to other common solvents such as acetonitrile (Cech and Enke 2001; Henriksen et al. 2005).

Despite this, some concern has been raised with regard to its reactivity. Since MeOH is a small alcohol, given properly acidic or basic conditions, it could potentially interact with polar functional groups such as carbonyls (Reusch 1999). In one case, Bateman *et al.* (2008) observed severe MeOH-induced artifacts in limonene SOA formed by esterification and acetal/hemiacetal reactions between MeOH and carbonyls (Bateman et al. 2008). McIntyre et al. reached similar conclusions with various humic substances (McIntyre et al. 2002). Other groups, such as Novotny *et al.* (2014), have shown that although these reactions exist to an extent, they do not significantly change the

composition of the sample. Although these reactions require sufficiently acidic conditions to occur, the reactions may occur within the sample itself. This could occur due to the large presence of carboxylic groups in complex organic matter, especially for oxidized aerosols (Heald et al. 2010; Ng et al. 2011; Zark et al. 2017). However, MLS is also rich in aromatic functional groups. Solvent interactions may be reduced in liquid smoke as the aromatic groups can stabilize adjacent carbonyls and reduce their reactivity.

The purpose of this investigation was to determine the extent of reaction artifact formation in mesquite liquid smoke (MLS) when using MeOH as a solvent. To do so, MLS was analyzed in both regular MeOH and MeOH-d₃ to refine the MF list and their potential reactive precursors. Reactions studied were based on those proposed in Bateman *et al.*, where precursors for esterification and acetal/hemiacetal reactions were located. This study has implications for the molecular characterization of complex organic mixtures with unknown analytes that may be susceptible to these reactions.

1.6 Thesis Structure

Each of the above-mentioned investigations are presented in separate chapters. The current chapter (Chapter 1) provides background information on the importance of complex mixtures, as well as the importance of method considerations used to measure and characterize them, such as the use of liquid smoke as a test sample. Chapter 2 details the specific materials and methods used to measure and analyze each of the aforementioned goals. This will include sample preparations and liquid smoke extraction

methods, instrumentation, and software parameters utilized for molecular formula assignment and data analysis.

Chapter 3 details an application of MS informatics, where hydrogen-deuterium exchanges in liquid smoke are observed and characterized. This chapter focuses less on artifact formation and more on how deep analysis of HDX could potentially offer insight into molecular functional groups or characteristics in mesquite liquid smoke sample components.

Chapter 4 investigates the possibility of creating reaction artifacts or electrospray adducts in liquid smoke samples with NH_4OH . Positive and negative ions in two different liquid smoke samples will be analyzed to study the likelihood and severity of NH_3 reactions or NH_4^+ adducts on certain molecular formulas.

Chapter 5 investigates the likelihood for methanol to interact with mesquite liquid smoke as a reaction artifact using analyses and methods similar to that of the ammonia/ammonium artifact. The general goal of this chapter is to determine which reactions are most likely with which MF, and how extensive these reactions are.

Chapter 6 describes the major conclusions and implications of each of the investigations in terms of sample preparation and analysis of complex organic mixtures. Potential ideas for future analyses will also be offered as the next steps to deepening the molecular understanding of these interactions.

2 Materials and Methods

2.1 Sample Preparation

Except where noted, all samples used in these studies were *Wright's* brand mesquite liquid smoke (MLS) or hickory liquid smoke (HLS), commercial smoke additives for food. They are water-concentrated products of wood-burning that are composed of thermally degraded lignin and cellulose derivatives (Montazeri et al. 2013, B&G Foods, Inc. 2019). Liquid smoke is an inexpensive surrogate for biomass burning samples useful for aerosol sample method development.

2.1.1 Hydrogen-Deuterium Exchange Sample Preparation

2.1.1.1 Acid Mix Sample Preparation

To understand the general mechanism for HDX in negative ESI, a mixture of four acids were prepared in both regular MeOH (control sample) and MeOH-d₁ (MeOH with a deuterated hydroxyl group). The standard compounds tested were vanillic acid (C₈H₈O₄ – one aromatic COOH group and one aromatic OH group), suberic acid (C₈H₁₄O₄ - two aliphatic COOH groups), citric acid (C₆H₈O₇ – three aliphatic COOH groups and one OH group), and polyacrylic acid (PAA, (C₃H₄O₂)_n with –COOH groups). Vanillic acid, suberic acid, and citric acid were selected based on both availability and structural variety. Specifically, vanillic acid is biomass burning relevant, and was used to investigate the impact of aromaticity on the rate of HDX. PAA was chosen to investigate the extent of HDX for polycarboxylic species. The “acid mix” solution was prepared by solvating each standard in 90:10 H₂O:ACN at a concentration of approximately 1000

ng/ μ L, sonicating as necessary. 1000 μ L of each standard dilution was then added to a single vial and mixed, then diluted with either regular MeOH or MeOH-d₁ to concentrations of approximately 100 ng/ μ L for each component to make them suitable for injection. Samples were placed in the freezer overnight until analysis could be completed.

2.1.1.2 Mesquite Liquid Smoke Sample Preparation

Two samples of MLS were prepared by filtering insoluble material and diluting it 100-fold into regular methanol and MeOH-d₁ to make concentrations suitable for the instrument without further dilution. Samples were stored in a freezer when not being analyzed.

2.1.2 NH₄OH Sample Preparation

Two samples of MLS and HLS each were prepared by diluting 1.0 mL of MLS or HLS respectively with 5 mL of LC-MS grade water. Both liquid smoke samples were a concentrated yellow-brown color. In addition, two lab blanks were made using 6 mL LC-MS water. All 6 jars were mixed on a shake table at 60 rpm for 2 hours to extract them. Each sample was filtered with polytetrafluoroethylene (PTFE) filters to remove insoluble material from the liquid smoke that may have interfered with ESI-MS analysis. All LS samples left some brown material in the filters, indicating a significant amount of insoluble material present in the samples. The filtrate from each of these samples was refrigerated overnight.

Each sample and blank was acidified with 54 μL of 1.0M HCl in preparation for SPE. In addition, solutions of NH_4OH were prepared. The regular NH_4OH solution was prepared by diluting 30% NH_4OH into MeOH in a 1:100 (v/v) ratio, while the $^{15}\text{NH}_4\text{OH}$ solution was prepared by diluting 3N ($\sim 10\%$) $^{15}\text{NH}_4\text{OH}$ into MeOH in a 3:100 (v/v) ratio.

SPE was then performed on each sample and blank using Oasis HLB SPE cartridges. The columns were conditioned with 3 mL MeOH and 3 mL H_2O (each LC-MS grade) and the first elution was performed with previously prepared 90:10 MeOH: H_2O for all samples and collected. The second elution was then performed using either regular or ^{15}N -labeled NH_4OH and collected separately.

Each sample elution was then blown down with N_2 gas to exchange the MeOH solvent to ACN (See Chapter 5), where samples were transferred to other vials and enough ACN was ultimately added to reach a final sample volume of 2 mL. Some brown residue was left on some sample vials, so these were rinsed with 200 μL of ACN and transferred again to vials before bringing the volume to 2 mL. Each sample and elution was then prepared for Orbitrap analysis by diluting 10-fold with ACN. The analysis was either done as is or with 1 μL (0.1%) of 90% formic acid (FA).

In addition, samples without SPE were prepared with MLS to investigate the effect of pH on the NH_4OH artifacts. These were prepared in a similar way to the SPE samples, except they were extracted directly in ACN, and the NH_4OH was added directly following this step in lieu of SPE to facilitate any adduct or reaction formation. Lab

blanks using water instead of MLS were also prepared. Each sample was then prepared for Orbitrap analysis by diluting 10-fold and adding 0%, 0.5%, or 0.1% FA.

2.1.3 Methanol Artifact Sample Preparation

Samples of MLS were prepared by filtering, then diluting MLS 100-fold into different methanol isotopes. For the purposes of this work, these will be denoted as MeOH or “control sample” (regular, LC-MS grade methanol) and MeOH-d₃ (methanol with the three deuterium isotopes on the methyl group). Samples were diluted at the instrument if necessary to make concentrations suitable for analysis and were placed in the freezer when not being analyzed.

2.2 Ultrahigh Resolution Orbitrap Elite Mass Spectrometry Analysis

All data collection was performed using an ultrahigh resolution Orbitrap Elite hybrid mass spectrometer (Thermo Scientific) at the Michigan Technological University Chemical Advanced Resolution Methods (ChARM) laboratory. The instrument was tuned with an external calibration solution for mass accuracy each day prior to analysis. Data was collected with a resolving power of 240,000 using either positive or negative electrospray ionization (ESI) within the Xcalibur software (Thermo Scientific). Samples were injected directly into the instrument with a 250 μ L syringe until a stable electrospray was achieved.

Mass spectra for the acid mix were obtained in negative ESI mode due to their acidity, making them highly susceptible to deprotonation. Positive mode was not studied.

Four total spectra were obtained. Two spectra for each solvent (Regular MeOH and MeOH-d₁) were collected, one with a scan range of m/z 100-800 and one with a scan range of m/z 100-2000. The higher scan range was obtained to analyze the HDX capabilities of various lengths of the PAA. 100 scans with spectrum averaging were collected using tune parameters that optimized the electrospray signal.

All liquid smoke mass spectra were studied using positive and negative ESI with optimized tuning parameters after reaching a stable electrospray. 200 scans over the range of m/z 100-800 were collected with spectral averaging.

2.3 Data Processing and Molecular Formula Assignment for Liquid Smoke Samples

For the HDX and MeOH artifact studies (Chapters 3 and 5), all liquid smoke sample masses were initially recalibrated using the Composer64 software (Sierra Analytics), using an internal set of recalibrants found in the liquid smoke samples. MF were tentatively assigned to check the average mass error (< 1 ppm) and minimum standard deviation. Recalibrated masses were exported as a mass list with a noise cut applied. This procedure was similar for the NH₄OH artifact study (Chapter 4) but recalibration was performed using a newly designed custom R-based software, MFAssignR (Schum et al. 2019). MFAssignR was used since it is capable of flexible recalibration and assigning molecular formulas with various isotopes, including deuterium and ¹⁵N. MFAssignR was also used for molecular formula assignment of all liquid smoke samples, where chemically unreasonable MF assignments are removed

through internal QA parameters (Schum et al. 2019). All samples were assigned with C, H, N, and O, with other heteroatoms dependent on sample and ionization mode.

The signal-to-noise ratio and full molecular formula ambiguity were considered important for molecular formula assignment. The signal-to-noise ratio was calculated for each sample using the KMDNoise function as described by Schum (2019). Full ambiguity is a parameter within MFAssignR that includes all chemically reasonable molecular formulas for a given exact mass peak without further prioritization. With the incorporation of isotopes into the molecular formula assignment, full ambiguity must be used as it is highly likely for a CH(D)O formula to be assigned as a CHNO formula or vice versa. This is due to the very small mass differences that can arise between the two. For example, it is difficult to know whether a formula should contain C₇ or O₅D₄, which have a mass difference of only 0.32 mDa, as shown in Figure 2.1. This does not occur as frequently with H due to the difference in mass defect between D and H. Therefore, it is appropriate to retain all of the chemically reasonable MF assignments in isotopically labeled samples and then manually filter those based on matching to the non-labeled control sample. The goal was to ensure that the molecular formula assignment within labeled solvents were as consistent as possible with the control solvents, since the compositions should be overall the same in both solvents with the exception of artifact products that include isotopes.

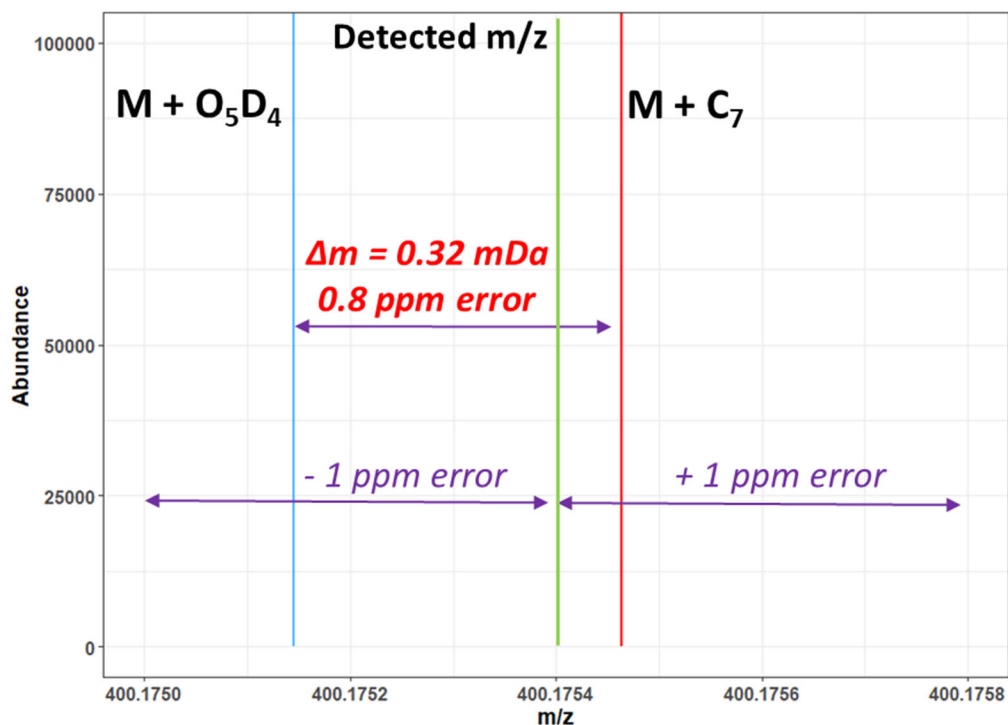


Figure 2.1. Mass error differences between O_5D_4 and C_7 -containing MF. “M” is the remainder of the molecular formula, constant between the two masses.

2.3.1 Molecular Formula Assignment Parameters

2.3.1.1 Hydrogen-Deuterium Exchange

Molecular formulas were not assigned to the analytes of the acid mix samples since the compositions were relatively simple, and only select major peaks were investigated, including the most abundant PAA ions.

The liquid smoke assignment parameters were defined differently for each ionization mode. The negative ions in regular MeOH were assigned using two passes: Pass 1 assigned MF without N and a 6*SN cut; Pass 2 assigned MF with 3N. MeOH- d_1 formulas were similarly assigned with two passes: Pass 1 assigned MF without N, up to 3

D, a 6*SN cut, and full ambiguity; Pass 2 was the same as Pass 1, but included up to 3 N. The positive ions in regular MeOH were assigned with up to 3 N and 1 Na, and a 6*SN cut. Positive ions in MeOH-d₁ were assigned similarly expect with up to 3D and full ambiguity. Only 3 D were assigned due to the small mass differences that can occur and complicate accurate assignment with 4 D.

Since there are substantially fewer CHNO compounds than CHO compounds expected in the negative ion mode, it is more likely that a CH(D)O formula should be assigned instead of a CHNO formula. Because of this, CHO compounds were assigned first, then CHNO was assigned to the remaining peaks. This approach was less justifiable for the positive mode ions since CHNO dominate. In both cases the molecular formula assignments with D were matched with CHO in the control since the compositions should be overall the same in both solvents, with the exception of HDX.

2.3.1.2 *NH₄OH Artifact*

MFAssignR was used to determine recalibrants for the NH₄OH study. The recalibrants were found within 3 ppm mass error in negative ESI, and 6 ppm mass error in positive ESI. High mass measurement errors in the mass range of m/z 100-250 were observed for the positive ESI files. Molecular formulas (MF) were then assigned using the MFAssignR. Negative ion MF were assigned with up to 3 N for all files and 3 ¹⁵N for isotopically-labelled samples. Positive ion MF were assigned with up to 5 N and 1 Na⁺ for all files, and 3 ¹⁵N for isotopically-labelled samples. 3 ¹⁵N were used instead of 5 to limit ambiguous MF assignments and reduce computational time. All MF were assigned after a 6*SN threshold was applied as determined using the KMDNoise function (Schum

et al. 2019). All MF were assigned with full ambiguity so the most likely MF could be selected manually. All assigned ions had an absolute mass error less than 3 ppm.

2.3.1.3 Methanol Artifact

A 6*SN cut was applied to all samples before formula assignment. The negative ions in regular MeOH were assigned with up to 3 N; negative MeOH-d₁ formulas were assigned similarly, but included up to 6 D and full ambiguity. The positive ions in regular MeOH were assigned with up to 5 N; positive MeOH-d₁ formulas were assigned similarly, but included up to 6 D, 1 Na⁺, and full ambiguity.

2.3.2 Reducing Ambiguity and Data Size

Accurate molecular formula assignment can be difficult when including multiple heteroatoms, especially alongside isotopes. For example, the more heteroatoms that are included, the more molecular formulas are possible for a given mass, especially at higher m/z values (Koch et al. 2007; Schum et al. 2019). In order to reduce this ambiguity, it is essential to ensure the assigned molecular formulas in isotopically labeled samples are comparative to others in the sample. The methods for doing so and reducing the size of the data sets vary by sample, solvent, and ion polarity. These methods are described in detail for each chapter.

However, one universal data filter was applied for all studies. Assigned MF in the isotopically-labeled sample were compared to those in the non-labeled (control) sample. Any formula that is assigned in the labeled sample, whether or not it contains a labeled

atom, should have an identical, non-labeled formula in the control to be a “valid” formula. This simplifies the mass list and removes any interferences from different solvent contaminants, and is the basis of this filter. The goal was to determine the species most likely to do so based on a few assumptions; hence, only a subset of the most probable molecular formulas are considered in these studies.

3 Characterization of Labile Hydrogen in Mesquite Liquid Smoke Using Hydrogen-Deuterium Exchange

The following sections discuss the results of the hydrogen-deuterium exchange (HDX) experiments with mesquite liquid smoke (MLS). First, the HDX capabilities of four acidic compounds will be investigated as a potential method for improving MF validity and reducing the data size in MLS. Then, the HDX capabilities of negative and positive ions are addressed separately, where less likely formulas are removed based on several justified data filters and molecular parameter trends for HDX-capable species are analyzed. Lastly, positive and negative HDX-capable ions are compared to understand similarities and differences between the two sets of ions.

3.1 Hydrogen-Deuterium Exchange of Acid Mix

The obtained spectra of the Acid mix are given in Figure 3.1. Notably, the low molecular weight standard compounds were in high abundance due to their high ionization efficiency and relatively higher concentration. Citric acid (m/z 193.03) is the most abundant peak in the spectrum and is approximately four times more intense than the most abundant PAA peak (m/z 473.17), due to the variable length of the PAA polymer. In both the regular MeOH and MeOH- d_1 mass spectra, evidence of multiply

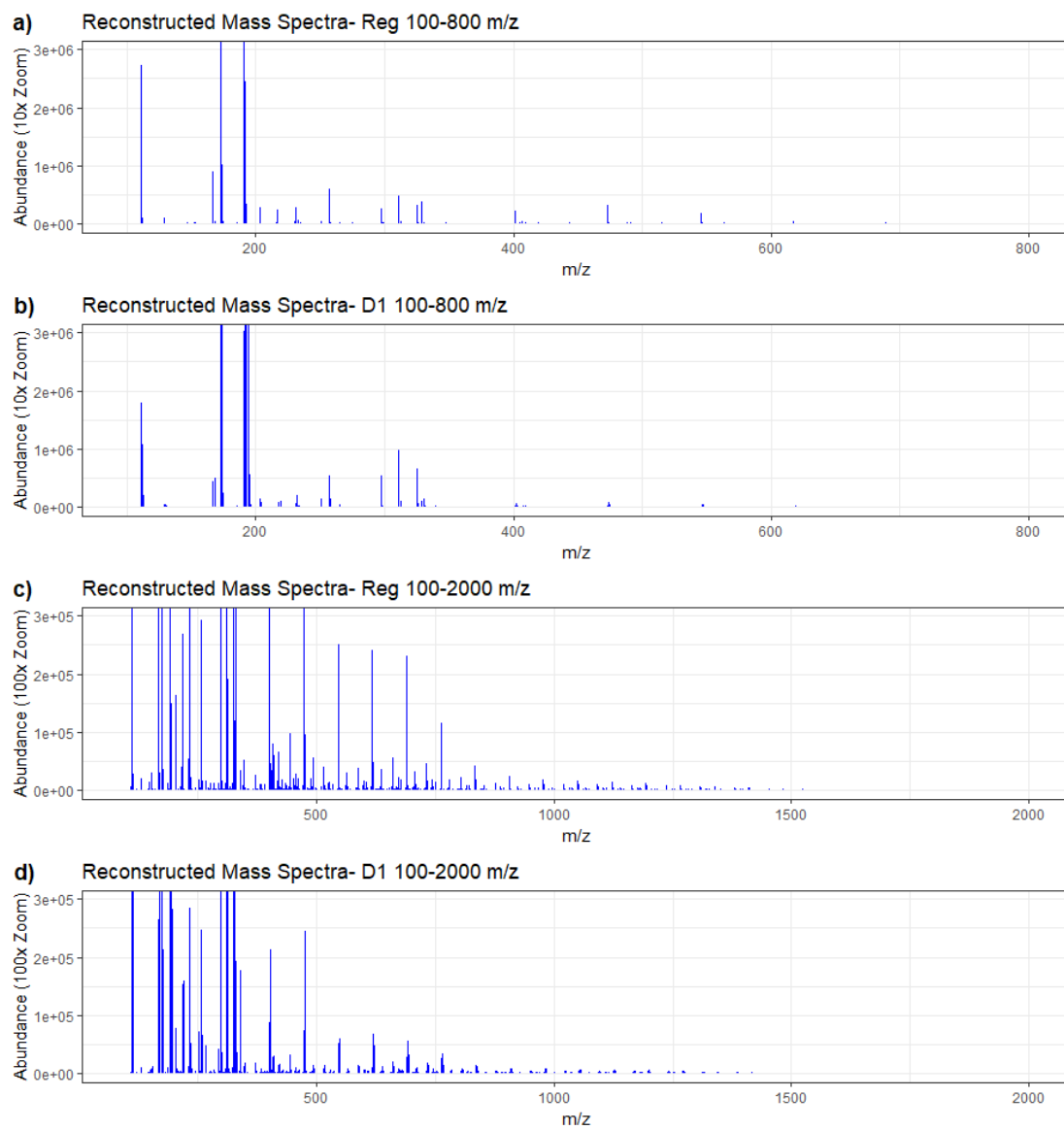


Figure 3.1. Reconstructed mass spectra for Acid Mix in a) Regular MeOH with m/z 100-800 scan range; b) MeOH-d₁ with m/z 100-800 scan range; c) Regular MeOH with m/z 100-2000 scan range; d) MeOH-d₁ with m/z 100-2000 scan range.

charged peaks are observed, but they were not the dominant peaks and thus were not considered for this analysis. For PAA, all major peaks were expected at mass intervals of m/z 72, indicating only singly charged species.

There is a noticeable increase in the ion complexity with the deuterated solvent. All of the standard compounds have multiple peaks associated with them, indicating strong evidence of HDX for each species. In particular, the higher molecular weight PAA polymers have a higher ion density, indicating higher levels of HDX.

Examination of each major species was done to investigate the nature of the HDX process, including all low-MW standards and several PAA peaks, as shown in Figure 3.2. Each subsequent high-intensity ion is an additional D substitution, and any low intensity ion in close proximity to the D peaks are ^{13}C signatures. In all low-MW standard species, HDX occurs up to the supposed maximum number of exchangeable protons, *i.e.*, 2 for suberic acid and 4 for citric acid. Ions with only some protons exchanged were also observed in all cases. However, the ions with all possible H exchanged for D are less than 5% the intensity of the base peak for each species. This is likely due to the ionization mechanism for negative ESI, which requires the analyte to be deprotonated to be detected. The most likely ionized functional group is the most acidic one, and thus, the one D that would be most labile and likely to be involved in HDX.

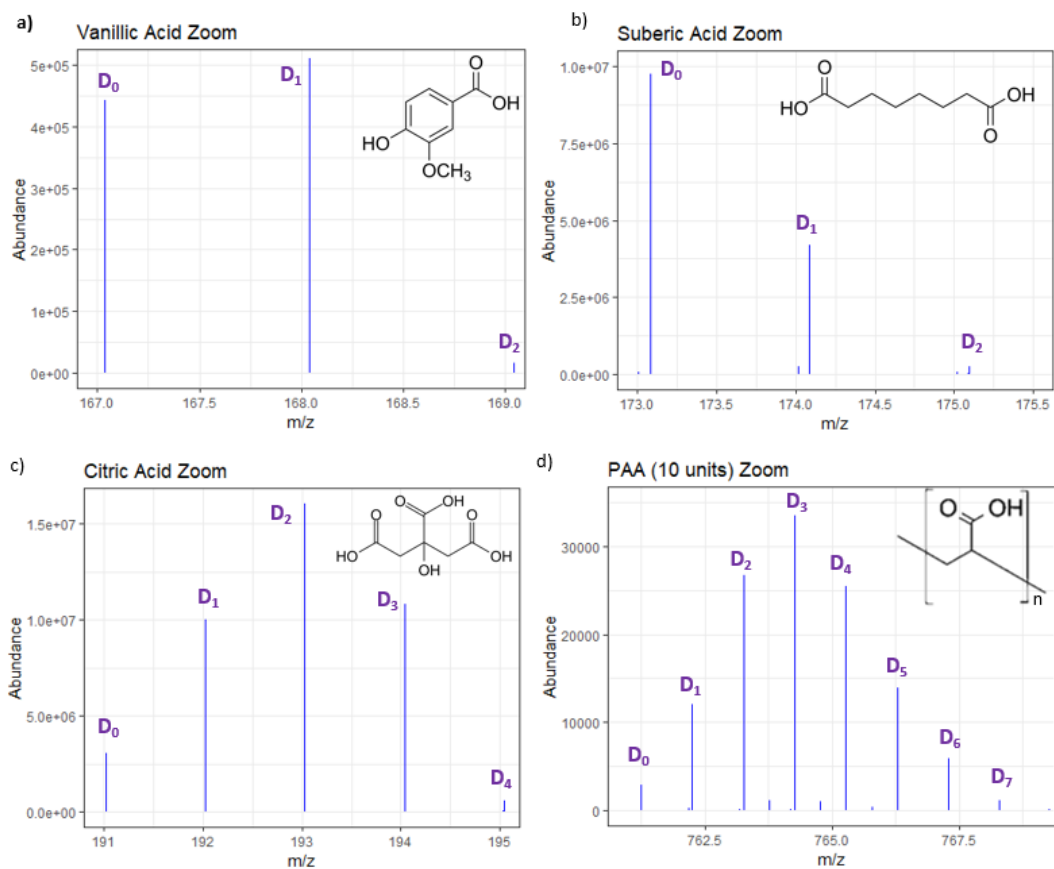


Figure 3.2. Magnification of Acid Mix spectra in MeOH- d_1 for (a) vanillic acid; b) suberic acid; c) citric acid; d) PAA with 10 acrylic units. Major peaks are labeled by number of deuterium present. Structures of each species are included as inserts in each spectrum.

Interestingly, each low-MW standard compound had a different ion distribution indicating the number of exchangeable H in each. Suberic acid, for example, had a roughly linear decay in the intensity of subsequent D substitutions, with regular suberic acid having the greatest intensity and suberic acid- d_1 having about 45% the D_0 intensity. Vanillic acid, on the other hand, had regular and D_1 - isotopes of approximately equal

intensity. This is potentially because the carboxylic proton is easier to exchange due to the resonance stabilization of the benzene ring.

Citric acid has all of the possible combinations of exchanges, with detected isotopes up to 4 exchanges. Unlike the previous two compounds, the relative abundance of each isotope follows a simple binomial Gaussian distribution, with D₂ having the greatest abundance and other isotopes having abundances tailing in an approximately 1:4:6:4:1 ratio. This implies that for citric acid, all labile protons have a 50% chance of being exchanged for D in MeOH-d₁ solvent. It could be hypothesized that this occurs due to the lability and/or number of carboxyl O-H bonds, as it does not occur with any of the other low molecular weight standards. For this multi-acidic molecule, the bond strengths for COOH and COOD are likely very similar since the acidic hydrogens are weak, so the detection of H and D substitutions are equally likely.

PAA shows a similar H/D exchange trend as observed for citric acid with each polymer length having a binomial distribution. However, the maximum number of exchanges is not the same as the maximum number of labile protons, or number of COOH groups in this case. For example, the PAA chain with m/z 761 is composed of 10 monomeric units and 10 COOH groups, however, the isotope of maximum m/z with a relative abundance greater than 1 is the D₅ formula, indicating there is an apparent limit to the extent of HDX. In these cases, however, the non-deuterated ion does not completely disappear until the polymer length is sufficiently above 1000 Da. Thus, it is reasonable to assume that the D₀ ion of a given species that can undergo HDX should exist in order to be considered valid for the purposes of the liquid smoke analysis.

3.2 Negative ESI Analysis of Mesquite Liquid Smoke

3.2.1 Preliminary Molecular Formula Composition

The ultrahigh resolution mass spectra obtained for mesquite liquid smoke (MLS) in negative ESI mode are given in Figure 3.3; Figure 3.3a contains the ions measured in regular MeOH, and Figure 3.3b contains the ions measured in MeOH-d₁. As expected, each mass spectrum is very complex due to the extensive composition of MLS. Each mass spectrum is separated by the elemental composition where CHO refers to molecular formulas containing only C, H, and O atoms, and CHNO refers to those containing only C, H, O, and N atoms. A few CHN and CH molecular formulas were assigned, but they were removed because they constituted <1% of assigned ions. Colors indicate the number of double bond equivalents (DBE) in the assigned molecular formula, with green representing CHO-only containing molecular formulas, blue representing CHNO-only containing formulas, and so on. An initial inspection of the MLS in MeOH mass spectrum indicates a high number of CHO compounds (~85% of assigned peaks), consistent with expectations for negative ESI analysis. Though scan ranges of m/z 100-800 were obtained for each sample, the signal begins to taper off drastically around m/z 350 for both regular MeOH and MeOH-d₁, though assignments exist up to m/z 700 for MeOH and m/z 650 for MeOH-d₁. Visually, the ions in MeOH-d₁ are denser than those in regular MeOH, providing evidence of HDX in the sample. This is further supported by the increase in the total number of measured ions in MeOH-d₁, with 2811 ions being observed compared to 2004 in regular MeOH. The tall CHNO ions around m/z 300 are S-containing contaminants present in MeOH and were removed in future processing steps.

These were removed from analysis during future filtering steps. Likewise, the tall group of peaks in MeOH-d₁ from m/z 250-300 are contaminants present in the blank solvent.

These were also removed in future filtering steps.

van Krevelen plots for MLS in each solvent are given in Figure 3.4, again depicting the predominance of CHO formulas, but also highlighting the presence of CHNO formulas. These CHNO formulas, however, are inconsistent and scattered throughout the VK plot, likely due to the MeOH-d₁ solvent being less pure (non-LC-MS grade) than regular MeOH, or due to incorrect assignments. The color scale indicates the DBE values up to 20, a range standard for various complex aqueous organic matter. There is a supposed decrease in DBE values in the MeOH-d₁ sample that is likely due to the overall lower signal; any high DBE values in regular MeOH had abundances < 60000 and likely fell beneath the S/N in MeOH-d₁; MeOH-d₁ did have a lower detected m/z range and overall lower intensity for high m/z peaks.

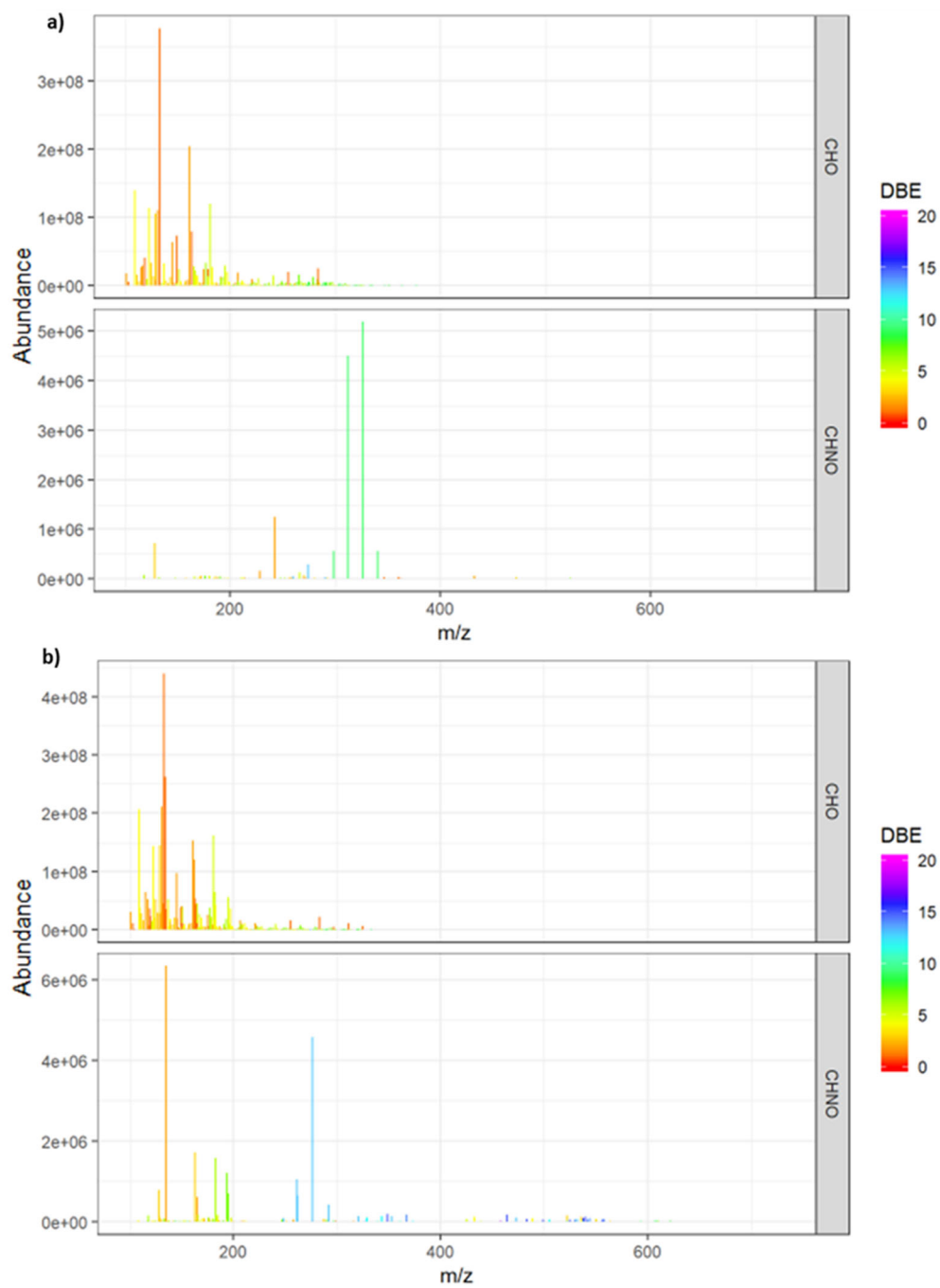


Figure 3.3. Reconstructed mass spectra obtained in negative ESI for (a) MLS in regular MeOH; (b) MLS in MeOH-d₁.

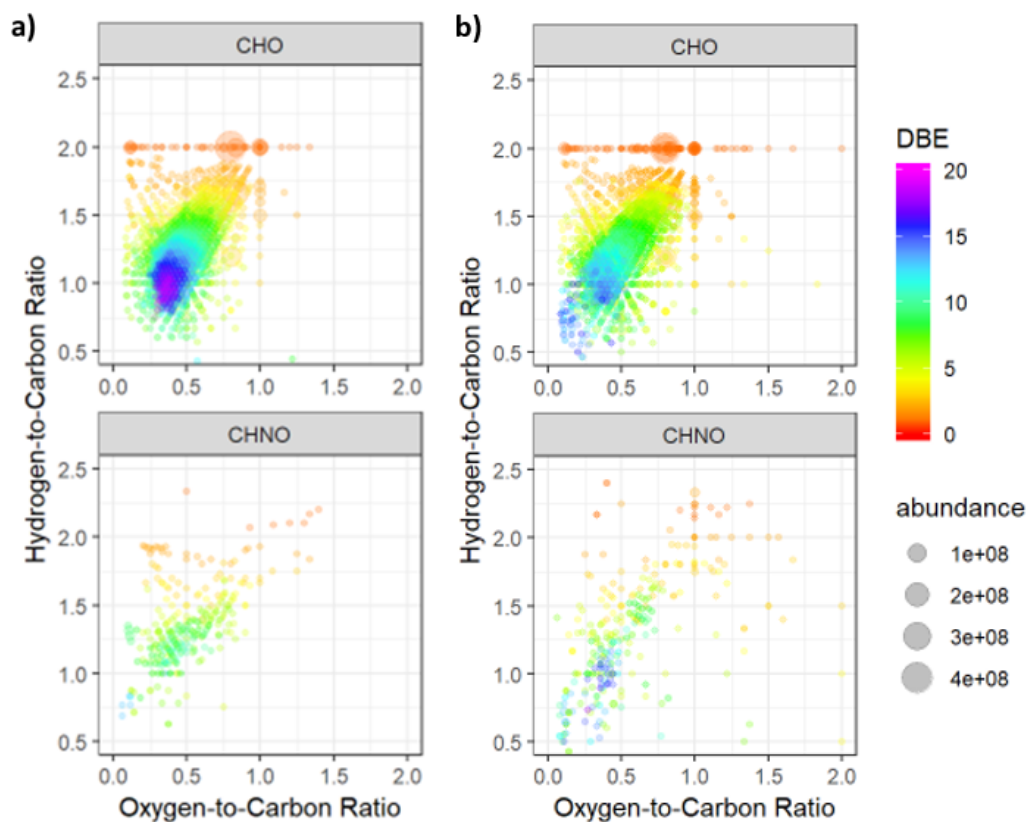


Figure 3.4. van Krevelen plots obtained in negative ESI for a) MLS in regular MeOH; b) MLS in MeOH-d₁. Points are sized by abundance and colored by DBE.

3.2.2 Filtering Formulas

Two data filters were applied to the MeOH-d₁ data to refine the data set and explore molecular characteristics of only the MF most likely to undergo HDX. One filter compared MF assignments to the regular MeOH, and the other used results from the acid mix to remove unlikely D-containing MF. The resulting VK plot after applying the first filter is given in Figure 3.5. The CHNO molecular formulas are less scattered, and the MeOH-d₁ molecular formulas are now more consistent with those of regular MeOH. The complexity of the CHO formulas is retained as expected.

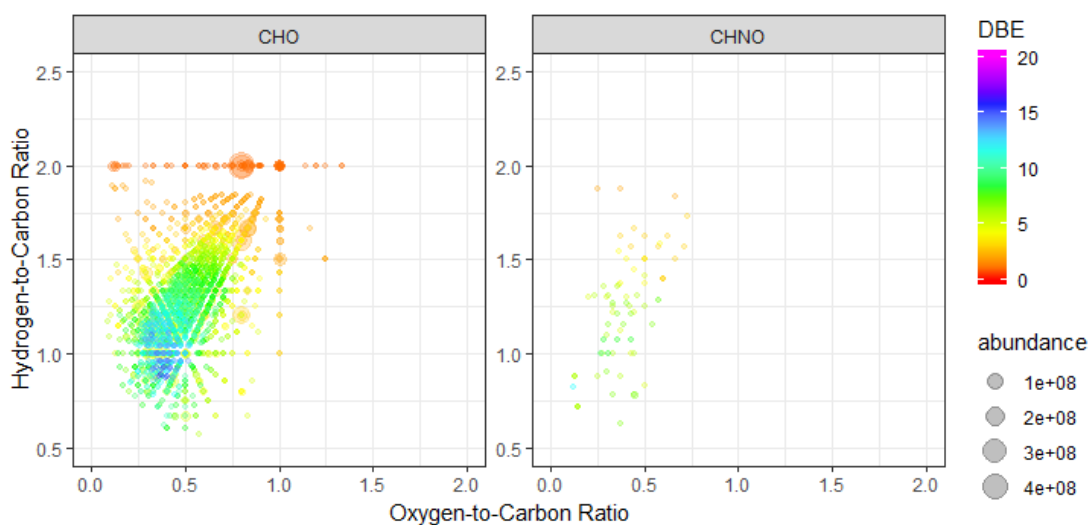


Figure 3.5. van Krevelen plot of negative ion MLS in MeOH-d₁ after applying first manual filter. Points are scaled by abundance and colors indicate DBE.

The reconstructed mass spectra resulting after the application of the first filter are given in Figure 3.6, where the mass spectra indicate the number of D atoms, and therefore, occurrences of HDX. The relative abundance scale has been zoomed to about 0.5% relative the tallest peak in the non-deuterated MFs to highlight the ion distribution patterns. D₁ ions follow a similar general abundance pattern to the non-deuterated formulas, with more intense peaks less than m/z 300 that taper off until about m/z 600. There is also a distinct dip in the assignments around m/z 400. D₂ and D₃ ions show two distinct groups of peaks instead of a continuous distribution, with a more sporadic and inconsistent pattern. The ion cluster separation could be a result of changes in the general structure of HDX-capable species from m/z 300-400, or it may be due to the MeOH contamination peaks which may have reduced the ion signal in the m/z 300-350 range. Notably, there are a few D₃ peaks that are more intense than their corresponding D₂

peaks, which is chemically unlikely due to the probabilistic nature of HDX. These ions were likely misassigned CHNO ions due to D₃ ions having masses in the nominal mass range of N. As a result, a second filter was needed to ensure the feasibility of these formulas.

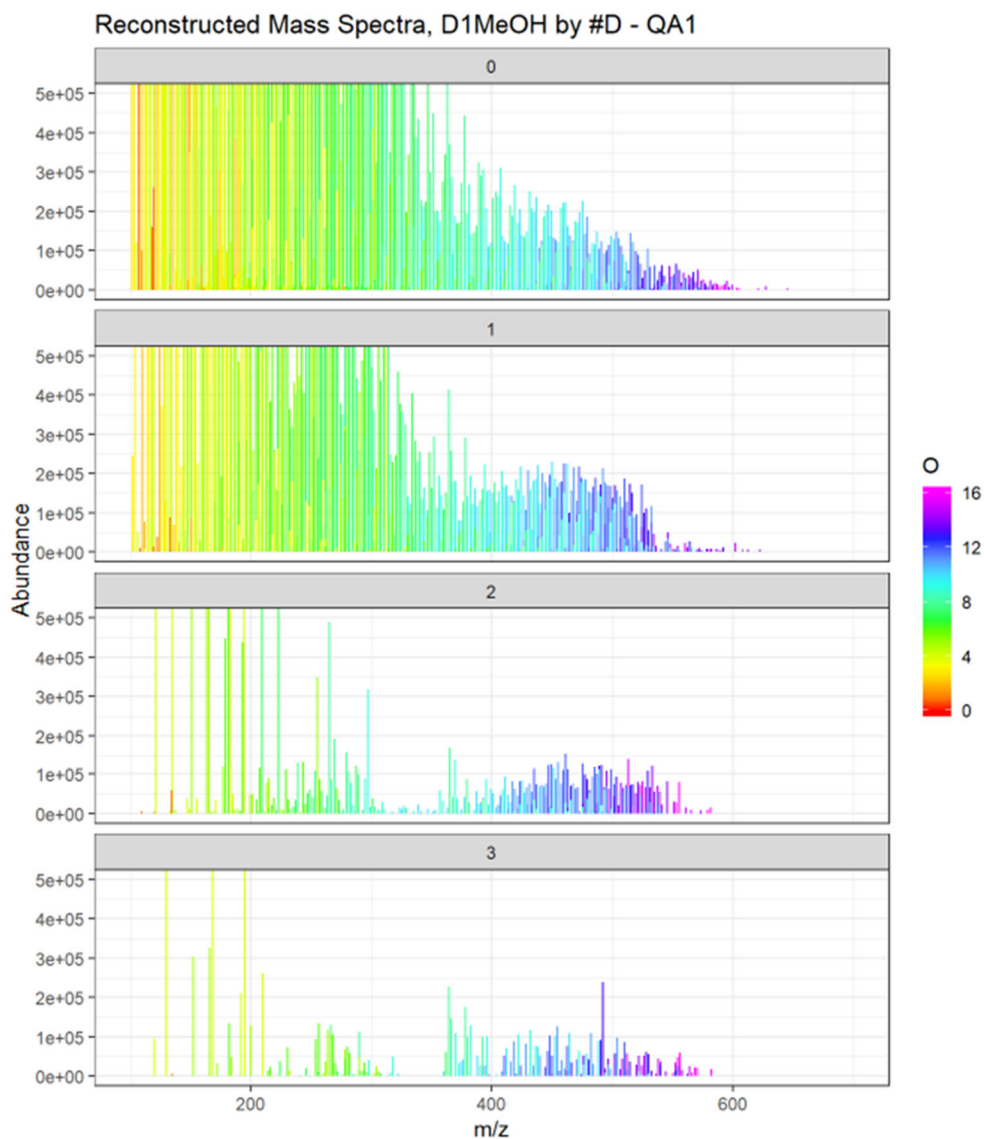


Figure 3.6. Reconstructed mass spectra of MLS in MeOH-d₁ (negative ESI) after applying first filter. Plots are separated by number of D and colors indicate number of oxygen.

The second filter applied to the HDX data is based on the probabilistic assumption that HDX is an equilibrium process, as indicated by the results from the acid mix. At any given moment, it is possible for an exchangeable proton to either be a proton (H) or a deuteron (D). Given this, for a molecule with multiple exchangeable protons, it seems unlikely that all of these protons will be exchanged for D without a lesser number of exchanges also being observed. Thus, for any assigned formula containing n D atoms, an assigned base formula must also exist with $n-1$ D atoms, $n-2$ D atoms, and so on. As an example, consider citric acid ($C_6H_8O_7$) from the acid mix standard. There are 4 exchangeable protons, but only a small percentage of citric acid molecules have all four simultaneously exchanged ($C_6H_4D_4O_7$). Compounds with 3, 2, 1, and no D were also detected, with the D_2 isotope being the most abundant. Therefore, it can be assumed that if the D_4 -citric acid formula and non-deuterated citric acid formula exist, then citric acid ions with 3, 2, and 1 D should also exist. This is the basis of this filter, where only the molecular formulas with a series of D atoms are considered “valid” formulas.

There are some limitations to this assumption. For one, it requires the D_0 peak to be above the SN ratio. As a result, any D-containing molecular formulas whose D_0 abundance is below the signal to noise are removed. In addition, some molecules may have variable rates of HDX due to variations in structural electronics, and therefore, some exchangeable protons will be more likely to exchange than others. This means there may be situations where some ions may or may not be present. However, the HDX rates of the model compounds, indicated that the D_0 peak did not disappear until about m/z 1000. Lastly, the electrospray ionization mechanism for negative mode deprotonates the

analyte, removing that hydrogen from the molecule. Therefore, a molecule that may ordinarily undergo HDX may not have an observable deuterated formula in negative ESI. It is also possible that there are multiple isomers within one ion peak, which have different capabilities of HDX.

The results after applying this filter are shown in Figure 3.7, where the plots are separated by the number of D in the formula. After these two filters were applied, no ambiguous molecular formulas were present, indicating the filters removed the ambiguity without manual manipulation. This filter also removed most CHNO compounds, to leave predominantly CHO compounds. This suggests N was not a significant factor in HDX rate for negative mode. The distribution of molecular formulas for each D was more ordered when the second filter was applied, and there is still evidence of two distinct groups of ions at lower and higher DBE, though it is more difficult to see in the non-deuterated formulas.

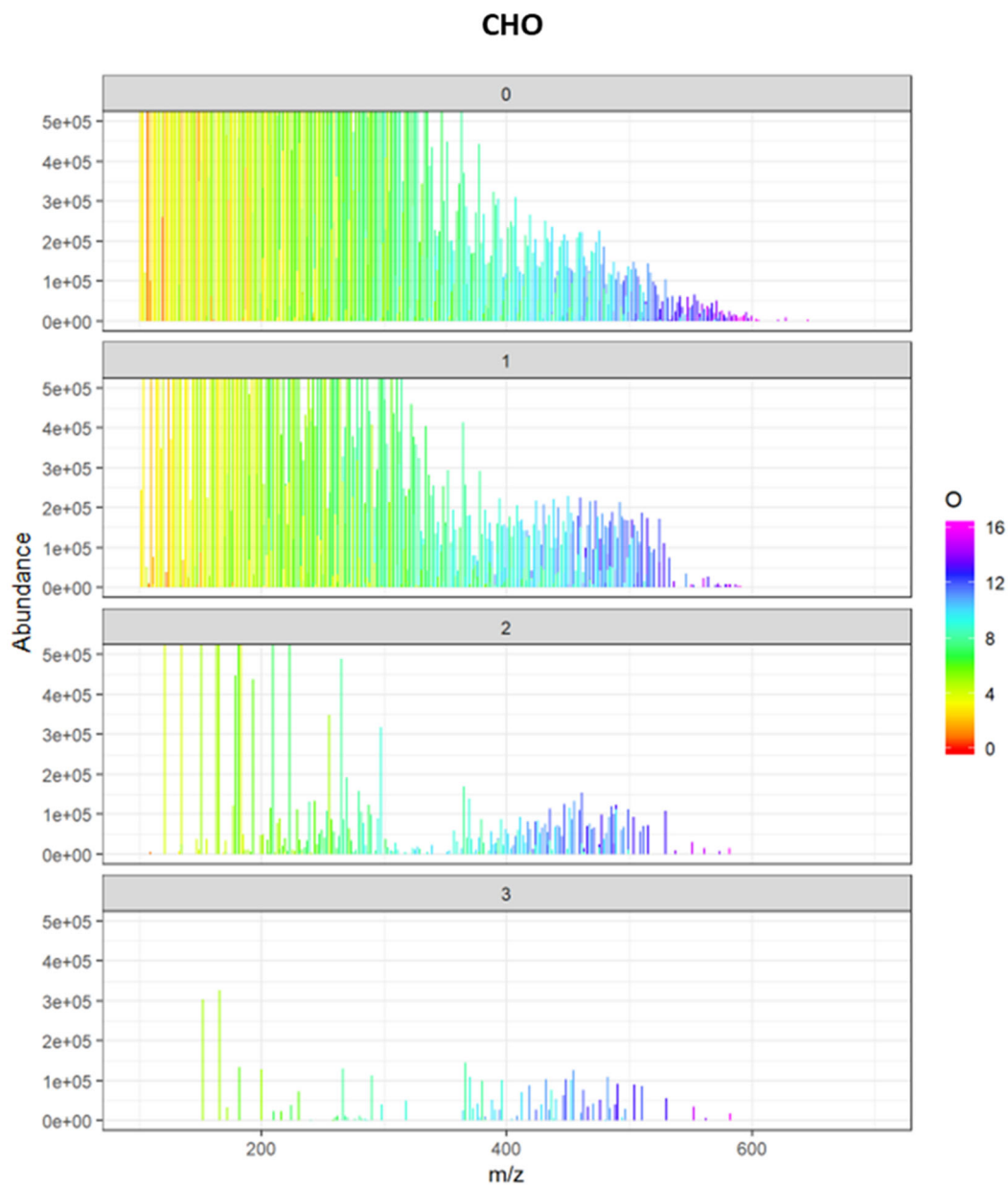


Figure 3.7. Reconstructed mass spectrum after final filtering of MLS d₁ data. a) CHO negative ions; b) CHNO negative ions. Spectra are separated by D and colors represent number of O.

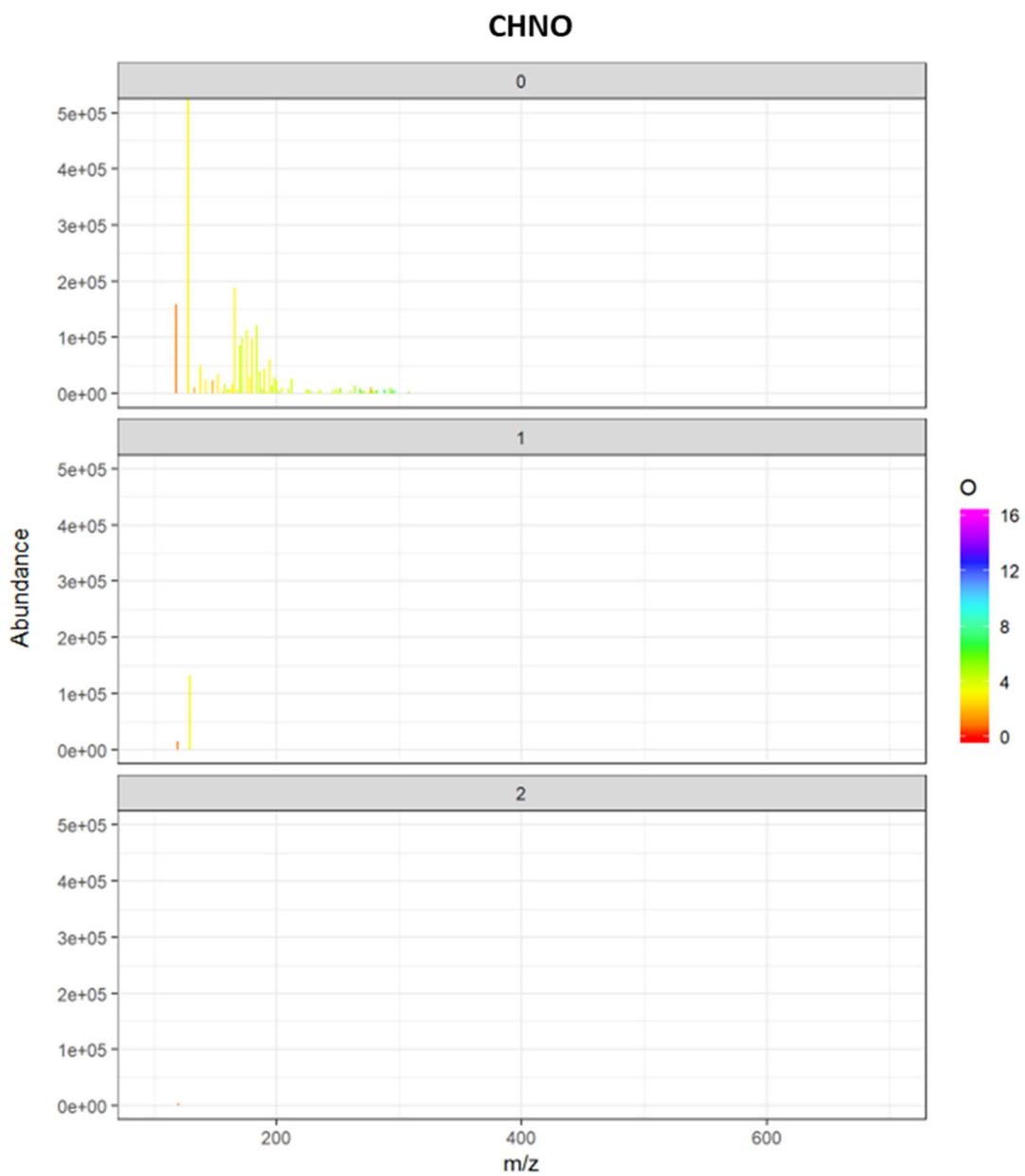


Figure 3.7 (cont.). Reconstructed mass spectrum after final filtering of MLS d_1 data. a) CHO negative ions; b) CHNO negative ions. Spectra are separated by D and colors represent number of O.

3.2.3 Trends in HDX-Capable Species

After the data filtering QA steps to isolate the most likely HDX MF, molecular trends for different numbers of exchanges were examined. The histograms for the number of oxygen atoms, DBEs, and carbon atoms of the HDX formulas are given in Figure 3.8(a-c). In these plots, the color scale represents the number of D atoms, with blue, cyan, orange, and red representing D₀, D₁, D₂, and D₃ species, respectively. In the case of D₀, both HDX capable and non-HDX capable formulas are included in order to compare trends to the overall sample. Since each molecular formula with a higher D substitution also has a lower D substitution, each larger D is a subset of a smaller D. Thus, the count on the y-axis is not a proportion, but an absolute count for each D. For example, there are approximately 100 species in the MeOH-d₁ sample that have 4 O atoms (Figure 3.8a), where 70 show evidence of 1 HDX and 10 show evidence of 2 HDX. There is a noticeable shift toward higher O numbers for and lower DBE for increasing numbers of D (Figure 3.8b). The increasing O count is expected; since N does not play a significant role in HDX in negative ESI, since any molecules capable of HDX should be more oxygenated. However, the decrease in DBE with increasing D seems to contradict this trend. One theory to explain this is that there are more molecules in MLS with multiple -OH groups than with multiple -COOH groups, with more saturated polyols being more available and capable of undergoing HDX. The carbon trends plot in Figure 3.8c depicts a bimodal distribution of molecular formulas emphasized in the mass spectral distributions in Figure 3.5, with peaks around 10 and 20 carbons for the D₂ and D₃ species. The bimodal distribution could be due to the influence of monoterpenes in biomass burning, as liquid smoke is a biomass burning surrogate (Montazeri et al.2012).

There is a minor shift towards fewer carbons as the number of H-D exchanges increases, potentially due to the propensity of lower carbon-containing species to have an increase in its O/C ratio, therefore being more likely to undergo multiple exchanges.

The averages of several molecular parameters were calculated for each HDX and are presented in Table 3.1. These were tabulated to determine the “average” molecular formula capable of undergoing certain numbers of HDX in MLS. The number of molecular formulas decreases exponentially for each additional HDX occurrence, which is partly due to the second data filter. As depicted in the values for the D₁ species, at least 63% of the total, non-deuterated molecular formulas in this sample have the capability of undergoing HDX, with 22% and 7.5% able to perform 2 and 3 exchanges, respectively. The values in this table quantify previous visual trends. The O and O/C values steadily increase with increasing D exchange, as does the H/C ratio. This increase corresponds to a decrease in DBE for each additional HDX, except for D₃ which is higher than D₂. This may be due to a shift to higher molecular weight compounds as the number of HDX increases (Figure 3.5), making the higher DBE species more dominant at higher exchanges. In all, the DBE is lower for formulas capable of HDX in comparison to the bulk average (D₀). This may indicate that HDX capability is dependent more on the number of OH groups as opposed to COOH groups in this sample. This can partially be seen by the average modified aromaticity index, or AI_{mod}, of each number of D.

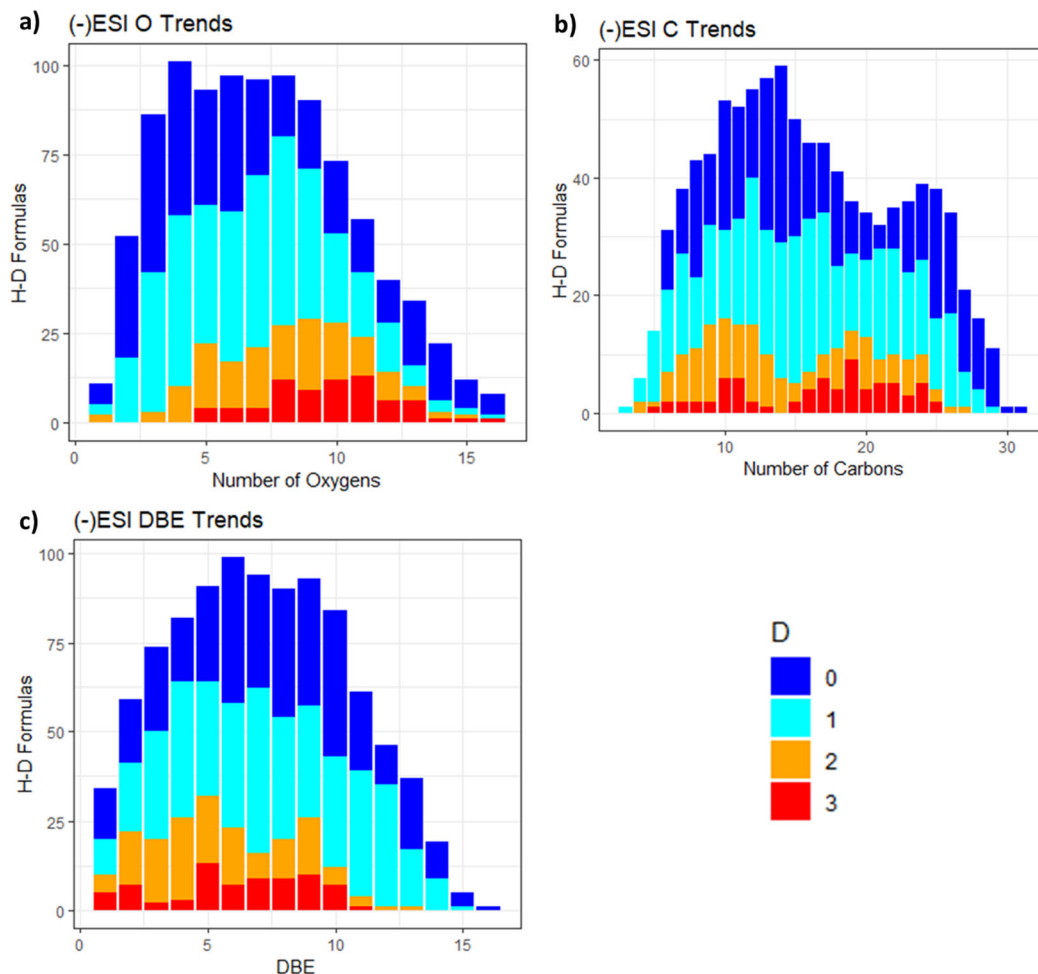


Figure 3.8. Various molecular formula trends with the number of deuterium for all formulas. Colors represent the number of deuterium. Plots represent number of D per (a) O; (b) DBE; (c) C.

AI_{mod} is a qualitative description of the aromaticity and unsaturated character of a MF based on the relative numbers of each atom proposed by Koch et al (2007). The calculation of this value is given in Equation 3.1, where C, H, N, and O represent the number of that particular element in the molecular formula. In the case of the deuterated samples, the number of D is included with the value of H. Values are then given an aromaticity descriptor, with the ranges given in Equation 3.1 as well. Based on the AI_{mod}

values, the average values for all number of D was an olefinic value. This implies D-containing species are rich in either carbonyl or alkene double bonds

$$AI_{mod} = \frac{\left(1 - C - \frac{O + H + N}{2}\right)}{C - \frac{O}{2} - N}, \text{ where}$$

$$AI_{mod} \geq 0.67 = \text{Condensed Aromatic}$$

$$0.5 < AI_{mod} < 0.67 = \text{Aromatic}$$

$$0 \leq AI_{mod} \leq 0.5 = \text{Olefinic}$$

$$AI_{mod} < 0 = \text{Aliphatic}$$

Equation 3.1. Calculation of modified aromaticity index, AI_{mod} , from Koch et al. (2007)

Table 3.1. Various parameter averages for MeOH- d_1 data in negative ESI.

D	Formulas	%Formulas	Mass	O	O_C
0	969	100.00%	325.95	7.2	0.47
1	614	63.40%	320.81	7.38	0.51
2	213	22.00%	334.13	8.53	0.61
3	73	7.50%	374.08	9.68	0.63
D	H_C	DBE	ZarkCOOH*	DBE-COOH**	AI_{mod}
0	1.27	7.09	3.57	3.52	Olefinic
1	1.28	6.83	3.73	3.11	Olefinic
2	1.41	5.69	4.55	1.13	Olefinic
3	1.42	6.16	5.27	0.9	Olefinic

*An estimate of a sample's overall COOH content, proposed by Zark et al. (2017).

**The difference between DBE and ZarkCOOH.

The COOH content of a bulk sample can be estimated as outlined in Zark et al. (2017). This estimate takes into account the H/C and O/C ratios of the species in the

sample and calculates the average number of COOH groups by scaling it to the maximum possible number of COOH groups in the sample. Equation 3.2 shows the calculation of the Zark et. al estimated COOH content for this sample, denoted as ZarkCOOH. The difference between the DBE and ZarkCOOH value is also given as DBE-COOH. The estimated COOH content increases with respect to HDX ability, up to an average of approximately 5 COOH groups for D₃ species. Conversely, the DBE-COOH value is decreasing with respect to the number of D, implying alkene functional groups may be playing a role in HDX at lower D.

$$ZarkCOOH = \frac{x_{min} - x_i}{x_{min} - x_{max}} [\#COOH_{max}]$$

Equation 3.2. Estimated COOH content of a sample, as described in Zark *et al.* (2017).

The [#COOH_{max}] term is the maximum number of COOH groups for the ion, or half the maximum O count rounded down. The x_i term is the O/H ratio of a particular formula, and x_{max} and x_{min} are the maximum and minimum O/H ratio in the sample. For this sample in negative ESI, this simplifies to

$$ZarkCOOH = \frac{0 - x_i}{0 - \frac{16}{42}} [\#COOH_{max}]$$

The van Krevelen plot for HDX capable species is given in Figure 3.9 with transparent dark blue points referring to the full MeOH-d₁ sample, and cyan, orange, and red corresponding to different numbers of HDX. The average O/C and H/C for each D (Table 3.1) are also plotted on top of the data with filled triangles, each representing the

average species for each number of D. Each average falls in a neat line of increasing O/C and H/C for increasing D.

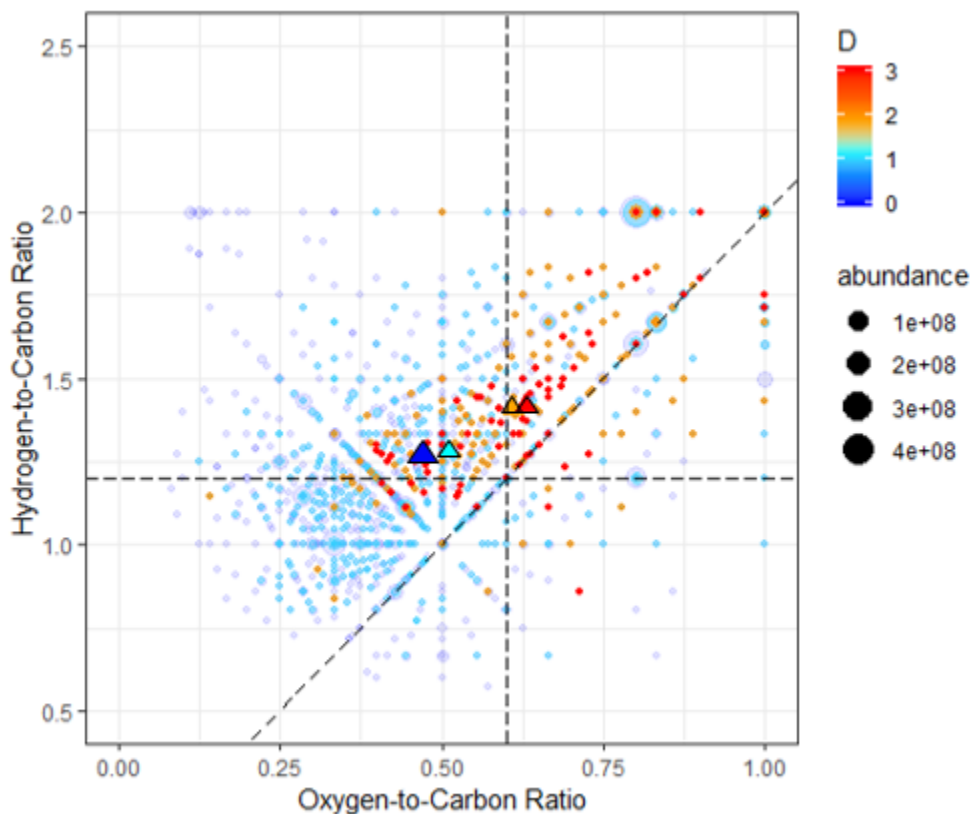


Figure 3.9. van Krevelen plot for filtered MLS data in MeOH-d₁ in negative ESI. Colors refer to the number of D in the formula, where D = 0 includes all non-deuterated molecular formulas with or without matches to D₁, and size corresponds to abundance. The colored triangles represent the average O/C and H/C for each D as given in Table 1.

3.3 Positive ESI Analysis of Mesquite Liquid Smoke

3.3.1 Preliminary Molecular Formula Composition

The positive ion ultrahigh resolution mass spectra for MLS are given in Figure 3.10, with Figure 3.10a measured in MeOH, and Figure 3.10b measured in MeOH-d₁.

Each plot is divided by the elemental groups. CH compounds were removed plot since they were so few and appeared to be less reasonable formulas due to their scattered nature. There were some positive ion CHN assigned that appeared reasonable due to distinct patterns in ion distribution, so they have been grouped together with CHNO ions for readability. As expected, there is a much larger number of CHNO ions than negative ions. The CHNO species range up to m/z 770 in both MeOH and MeOH- d_1 solvents. The ranges are significantly higher than those in negative ESI. In general, the positive CHO species had intensities approximately two times higher in MeOH- d_1 than MeOH. There is also a notable number of positive ion CHN, implying many amino functional groups were detected. In positive mode, CHNO compounds are the dominant species in MLS despite the high number of CHO compounds, with 59% N-containing molecular formulas in MeOH. In total, 5539 unique peaks were detected in MeOH and 7111 unique peaks in MeOH- d_1 , and 2165 of these peaks were ambiguous molecular formulas. This high number of ambiguous molecular formulas due to the high number of heteroatoms required for the formula assignment (N, D, and Na). These ambiguous formulas were reduced in several data filtering steps described below.

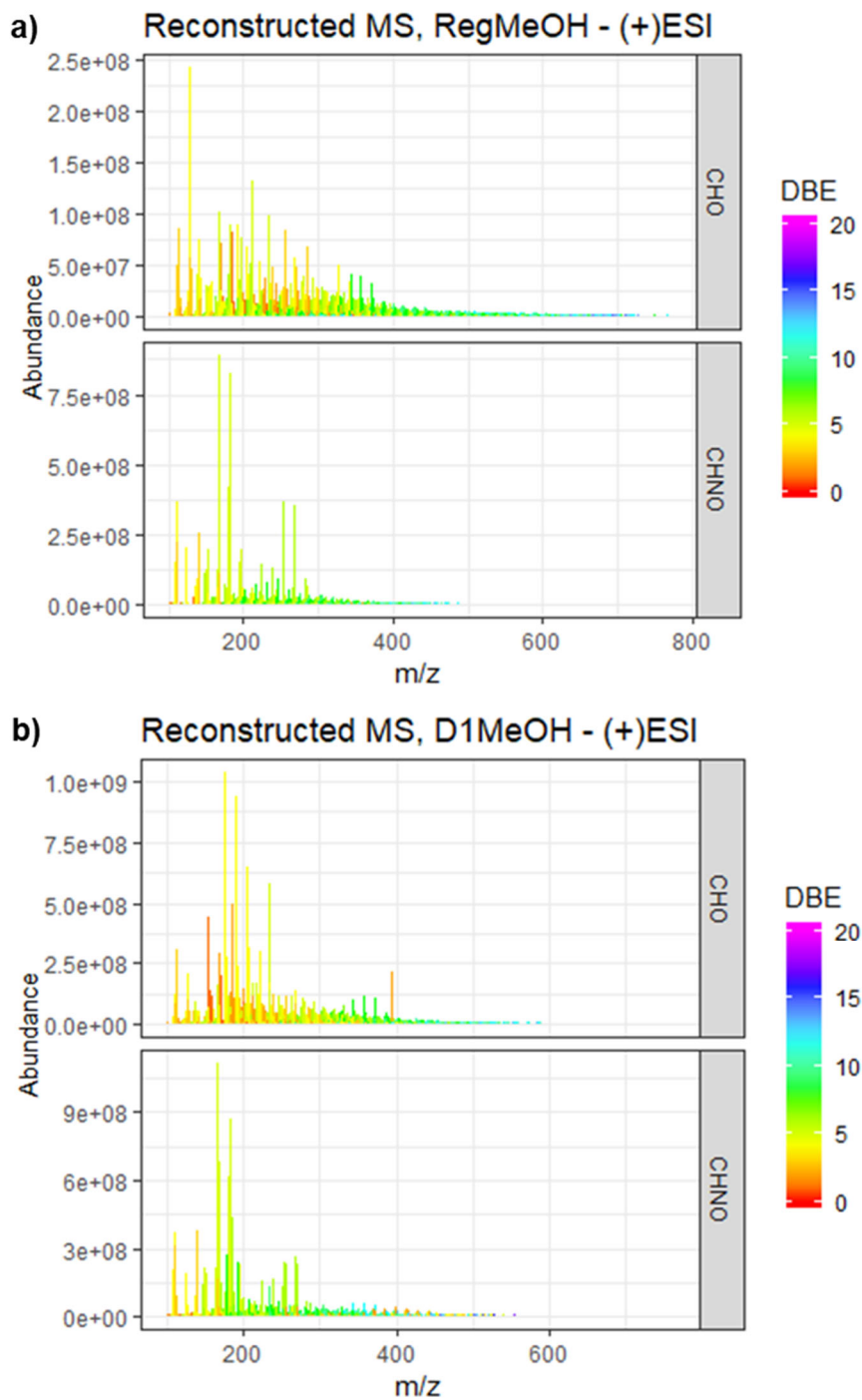


Figure 3.10. Reconstructed mass spectra obtained in positive ESI for (a) MLS in regular MeOH; (b) MLS in MeOH-d₁. Colors indicate the DBE of the assigned formula.

The van Krevelen plots for MLS in each solvent are given in Figure 3.11 separated by elemental group. Most of the compounds are in the low O/C and low H/C space with high DBE values, indicating the presence of highly unsaturated or aromatic species. As shown earlier, in Figure 3.10, there is a much larger presence of positive ion CHNO species than observed in the negative mode.

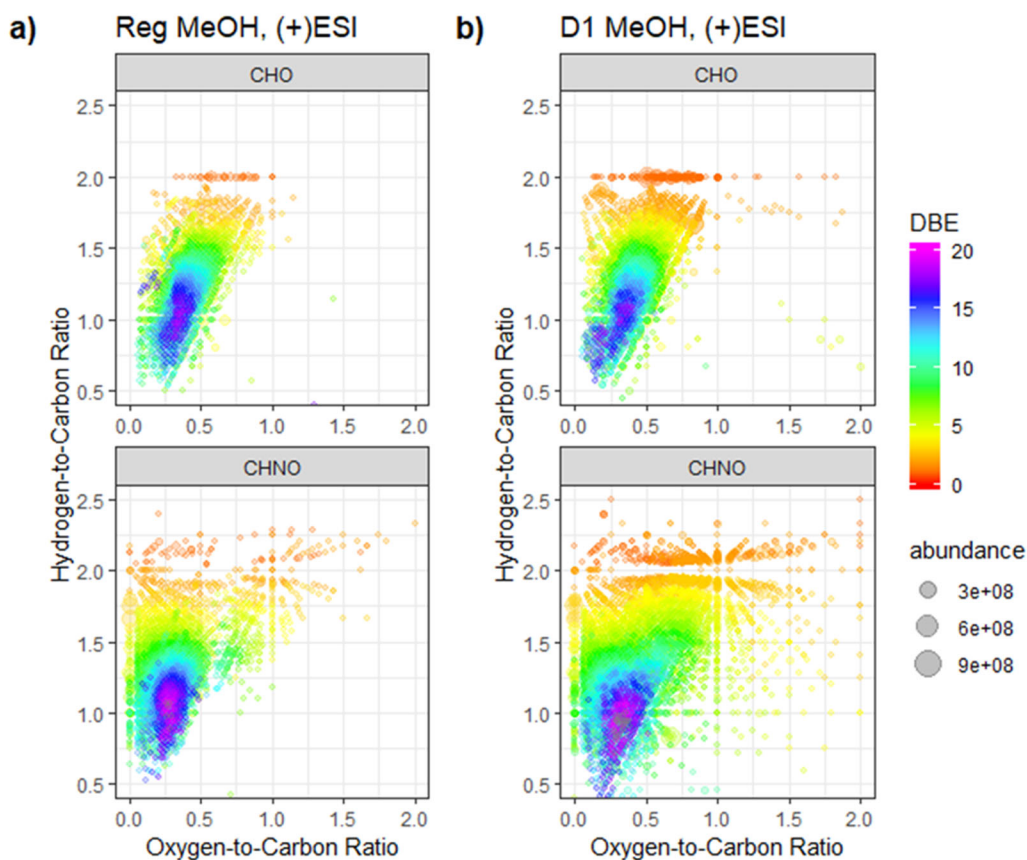


Figure 3.11. van Krevelen plots obtained in positive ESI for a) MLS in regular MeOH; b) MLS in MeOH-d₁. Points are sized by abundance and colors indicate DBE.

The range of elemental ratios of the positive ion species in MeOH-d₁ match those of MeOH, although there are several species outside the ranges. For example, the cluster of low H/C CHNO ions in MeOH-d₁ was not present in MeOH. Many of these MF are ambiguous, (Figure 3.12a), and will be removed in future data filtering steps. Most of the high DBE compounds are also ambiguous, although many of these will be matched to compositions in the MeOH.

3.3.2 Filtering Formulas

It is known (Kruve et al. 2013; Kruve and Kaupmees 2017) that there are multiple cations that can be present as adducts during positive mode ESI. Thus, it is desired to separate molecular formulas based on their adduct ion. The most frequent adducts are H⁺ and Na⁺ and provides a basis for separating the MF based on the number of sodium in the formula. This helps limit molecular formula assignments further. For a molecular formula to be considered valid in the MeOH-d₁ solvent, there must be a matching base formula in MeOH and a matching adduct, although it is possible for some MF to have both H⁺ and Na⁺ adducts (Kruve and Kaupmees 2017). After this first filtering step, there were a total of 3039 molecular formulas with H⁺ adducts and 682 molecular formulas with Na⁺ adducts, although some ambiguity was still present among these formulas.

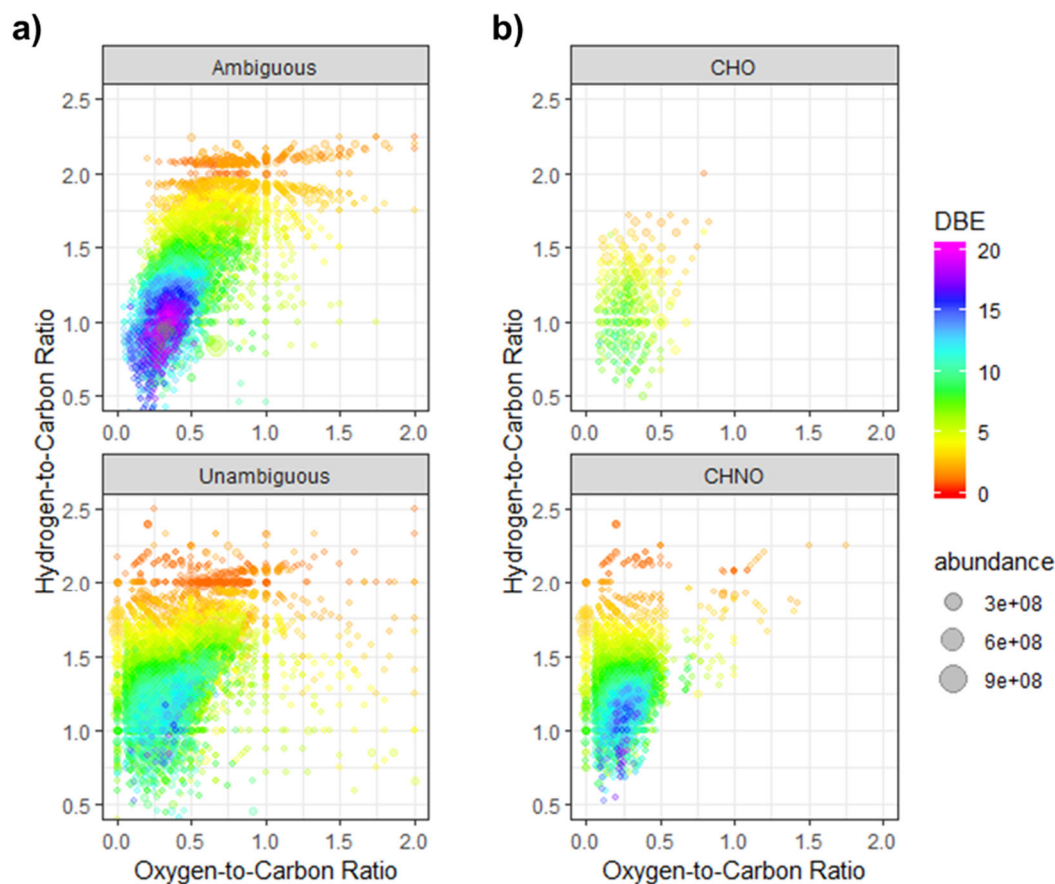


Figure 3.12. van Krevelen plots for MLS in MeOH- d_1 obtained in positive ESI before and after filtering. Symbols are scaled by abundance and colored by DBE. a) Ambiguous and unambiguous MF assignments before filtering; b) Post-filter MF assignments, separated by their molecular group.

The filtered results are shown in van Krevelen plots in Figure 3.12b, consisting of only H^+ adducts. This first filter alone greatly reduced the complexity of the data, removing many (though not all) scattered CHNO formulas and leaving a lower number of CHO compounds as indicated in the van Krevelen plot in Figure 3.12b. The reduced CHO content is reasonable due to the high N content of MLS, and highlights the need to use positive ESI to understand the impact of N on this analysis. In addition, (Kruve et al 2013) indicate that there is a tendency for H^+ adducts to be favored by CHNO

compounds and for Na⁺ adducts to be favored by CHO compounds. There may also be some interactions between the deuterated solvent and the mesquite liquid smoke that vary from those in regular solvent; however, this would require more investigation and is beyond the scope of this work.

Reconstructed mass spectra (Figure 3.13) for each number of D show the general abundance patterns of deuterium substituted species. The color scale indicates the number of O in the formula, which increases with mass. The ion abundance distributions in D₁ closely matches those of D₀ and D₂. However, several D₃ species are taller than their analogous D₂ formula. As before, this is unlikely based on results of the acid mix analysis (Figure 3.2).

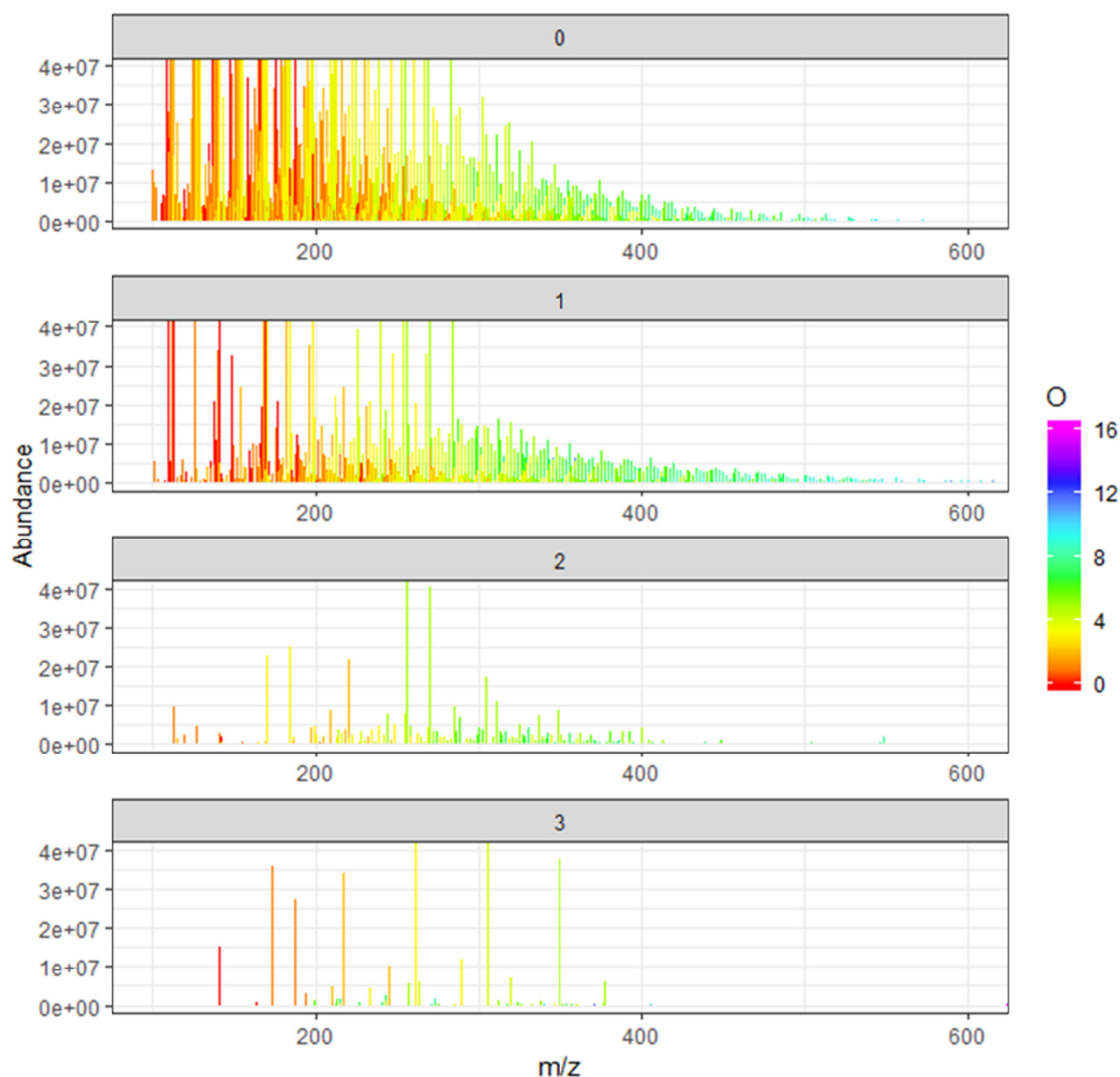


Figure 3.13. Reconstructed mass spectra of MeOH-d₁ H⁺ adducts of MLS after applying first data filter. Plots are separated by number of D. Colors indicate number of oxygen in the formula.

The data for the Na⁺ adducted species (Figure 3.14) show some notable differences from those with H⁺. As previously mentioned, there is a tendency for Na⁺ adducts to favor CHO compounds, and thus, there is a greater number of CHO ions than CHNO ions in the group-separated van Krevelen plots shown in Figure 3.14a. The

CHNO MF appear to be more scattered, which is likely due to their low number. Notably, the DBE of Na⁺ adducts is low.

The reconstructed mass spectra for the Na⁺ adducts of each number of D is given in Figure 3.14b. The peaks are colored by O number. D₁ and D₂ generally follow the general patterns present in the full data set (D₀), while D₃ has a few peaks that are more abundant than their analogous D₂ peaks. In addition, there is a higher number of O in the positive ion molecular formulas. This is seen in the reconstructed mass spectra, where O numbers are generally higher for each D in comparison to those detected in the negative ion mode.

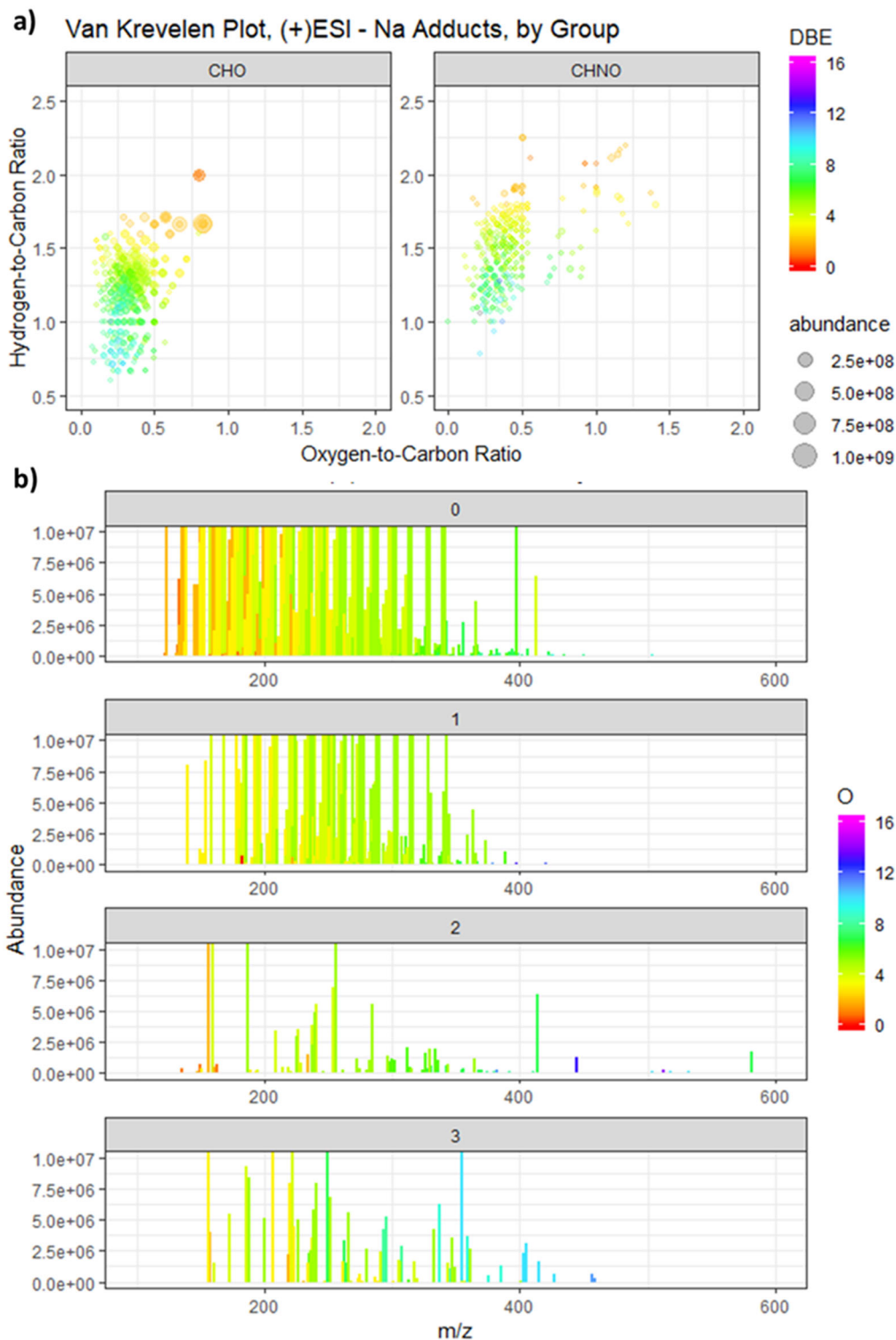


Figure 3.14. H^+ adducts of MLS in MeOH- d_1 after applying first manual filter. a) van Krevelen plot separated by group. Points are sized by abundance and colored by DBE; b) Reconstructed mass spectra separated by number of D. Peaks are colored by number of oxygen in the formula.

In the second filter, formulas were retained only if each previous D substitution was present. This narrowed the list of molecular formulas those most likely to undergo an exchange. This filter removed many of the ambiguous formulas, leaving only 88 ambiguous MF. All but one of these ambiguous pairs had one of two differences: C₃ND (mid-DBE) vs. H₄O₃ (low DBE) or C₂D₂ (mid-DBE) vs. H₆Na (low DBE). Since these MF made it through previous QA steps, it is not reasonable to remove formulas with more assigned deuterium, as they could be correct. In this case, the higher DBE was retained. This is based on the van Krevelen plot in Figure 3.12, which shows a vacancy of high DBE unambiguous formulas. These high DBE formulas were readily assigned in the regular MeOH. In addition, the higher DBE formulas had an overall lower error, and contained either a N atom or a H⁺ adduct, which are more consistent with the ambiguous MF.

The reconstructed mass spectra for H⁺ adducts from this filter are given in Figure 3.15, separated by number of D in the molecular formula. All H⁺ adducts that were capable of HDX in positive ESI were CHN or CHNO compounds, so CHO MF are not included in the spectra. There are noticeable patterns in the peak distributions up to D₃, where the few peaks present are more scattered. The DBE values for HDX-capable species is also relatively low.

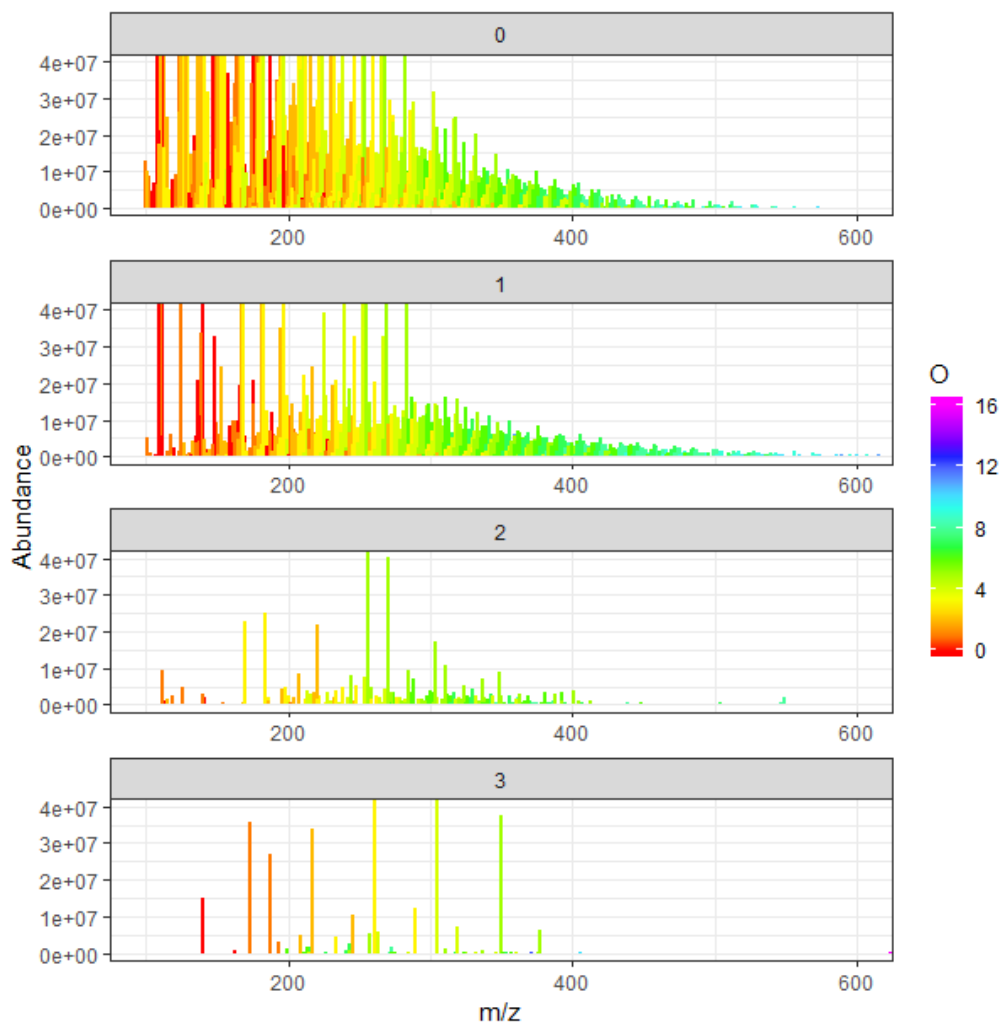


Figure 3.15. Reconstructed mass spectrum of N-containing H^+ adducts after final filtering of MLS MeOH- d_1 data. Spectra are separated by D and peaks are colored by number of O in the formula.

3.3.2.1 Trends in HDX-Capable Species

The histograms for O, N, DBE, and C trends of H^+ adducts are given in Figures 16(a-d) with blue, cyan, orange, and red indicating the D_0 , D_1 , D_2 , and D_3 species, respectively. A large proportion of the positive ions in MLS have at least one detectable instance of HDX, but there are very few D_3 formulas. The number of O in H^+ adducts

(Figure 3.16a) is heavily right-skewed for D₀ and D₁, with a few high O formulas being assigned in D₀. D₂ and D₃ have a more normal distribution centered around 5 O.

The N trends plot in Figure 3.16b depicts most of the HDX formulas as being N₁ compounds; almost all D₂ and D₃ species contain only one N atom. In addition, about 75% of the molecular formulas with 1 or 2 N were detected with HDX. Conversely, there are no instances of HDX for 0 and 3 N. The DBE distribution of H⁺ adducts in Figure 3.16c are approximately normally distributed around 6-8 DBE and seems to shift to lower DBE values for higher H/D substitutions. Similarly, the carbon trends in Figure 3.16d are also approximately normally distributed centered around 15-16 C for the full data set (D₀). D₁ is slightly shifted to higher C numbers, while D₂ and D₃ shift back towards lower values. There are some ions at higher O and C numbers that are likely incorrect MF assignments, though they were not ambiguous formulas.

Averages of molecular formula parameters are given for H⁺ adducts in Table 2. Of approximately 1500 H⁺ adducts observed in positive ion mode, over half of them exhibited evidence of HDX. D₁ and D₂ had a significant number of species, but less than 1% of formulas were capable of 3 exchanges. On average, the DBE and average mass decreased for increasing D substitution. In addition, the O number and O/C ratio increased with increasing D, as did the ZarkCOOH number. The estimated number of carboxyl groups for a given formula was calculated via Equation 3.2.

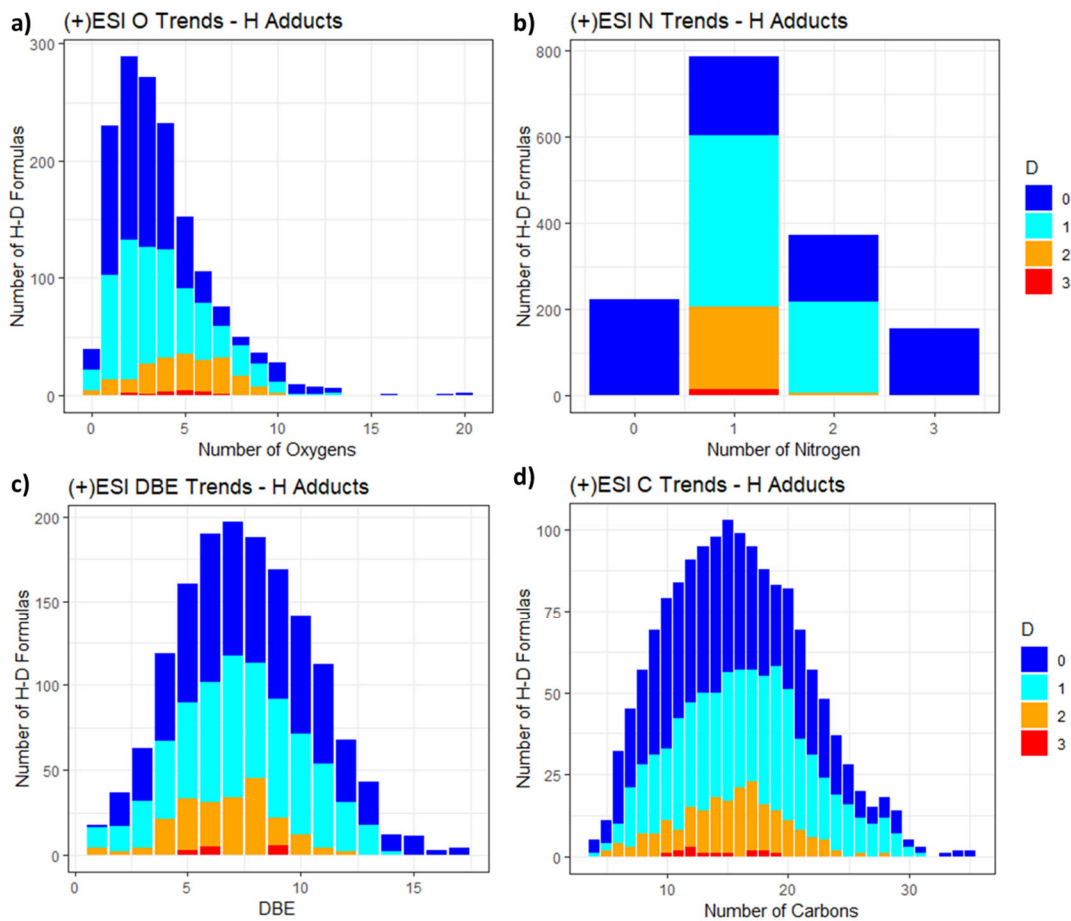


Figure 3.16. Various molecular formula trends with the number of deuterium for all formulas. Colors are based on the number of deuterium. Plots represent number of D per (a) oxygen; (b) DBE; (c) carbon.

Table 3.2. Various parameter averages for H⁺ adducted MF in positive ESI.

D	Formulas	%Formulas	Mass	O	N	O_C
0	1535	100.00%	289.46	3.79	1.3	0.24
1	822	53.60%	300.02	4.05	1.27	0.24
2	214	13.90%	296.82	4.9	1.03	0.31
3	14	0.90%	276.12	4.57	1	0.34
D	H_C	DBE	ZarkCOOH	DBE-COOH	Al _{mod}	
0	1.27	7.54	0.9	6.64	Olefinic	
1	1.33	7.26	0.9	6.36	Olefinic	
2	1.33	6.72	1.27	5.45	Olefinic	
3	1.22	7.07	1.3	5.77	Olefinic	

*An estimate of a sample's overall COOH content, proposed by Zark et al. (2017).

**The difference between DBE and ZarkCOOH.

The van Krevelen plot in Figure 3.17 indicates the relative location of each D species for H adducts. Colors indicate the number of D in the MF, and the symbol size indicates the abundance of each peak. The average O/C and H/C for each number of D are denoted with triangles of the respective colors. Notably, there is a shift toward higher O/C ratio for increasing number of D, though there is little trend in H/C.

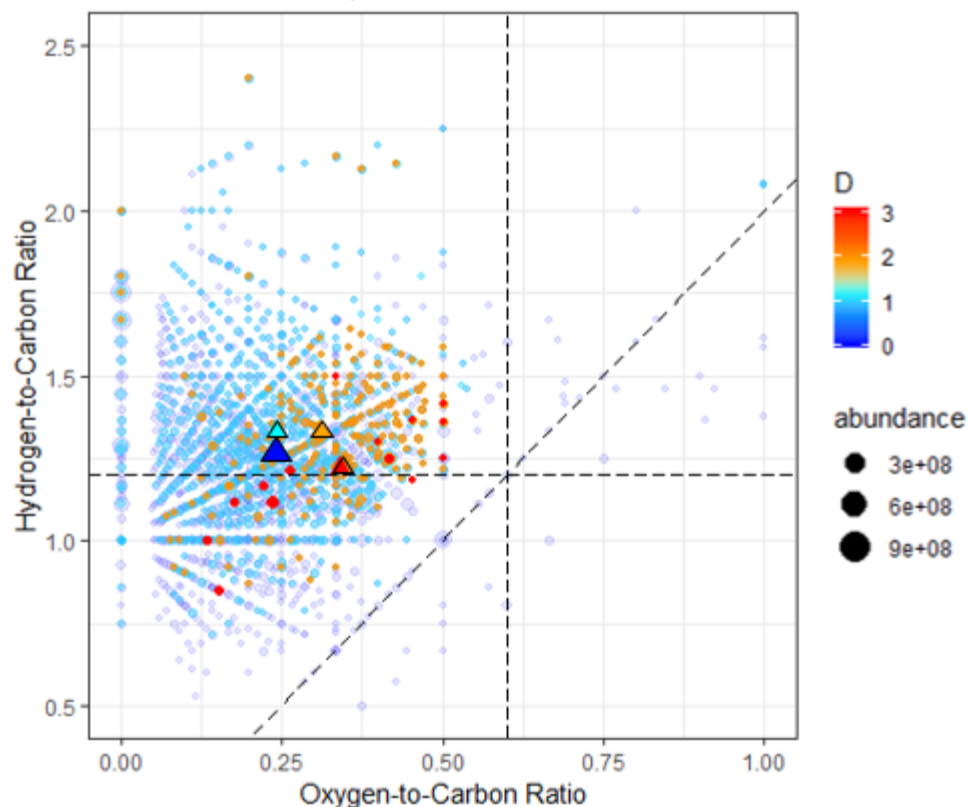


Figure 3.17. Van Krevelen plot for H adducts of filtered MLS data in positive ESI. Colors refer to the number of D in the formula, and size corresponds to abundance. The colored triangles are the average O/C and H/C for each D as given in Table 2.

The reconstructed mass spectra for Na^+ adducts after data filtering are given in Figure 3.18, again separated by number of D and the elemental group. In stark contrast to the MF with H^+ adducts, those capable of HDX with Na^+ adducts are dominated by CHO compositions. Ion distributions of D_1 and D_2 match closely to the overall data set (D_0), while there are very few formulas that display 3 exchanges. There appears to be only a small variation in the number of O, compared to a wider variation for H^+ adducts.

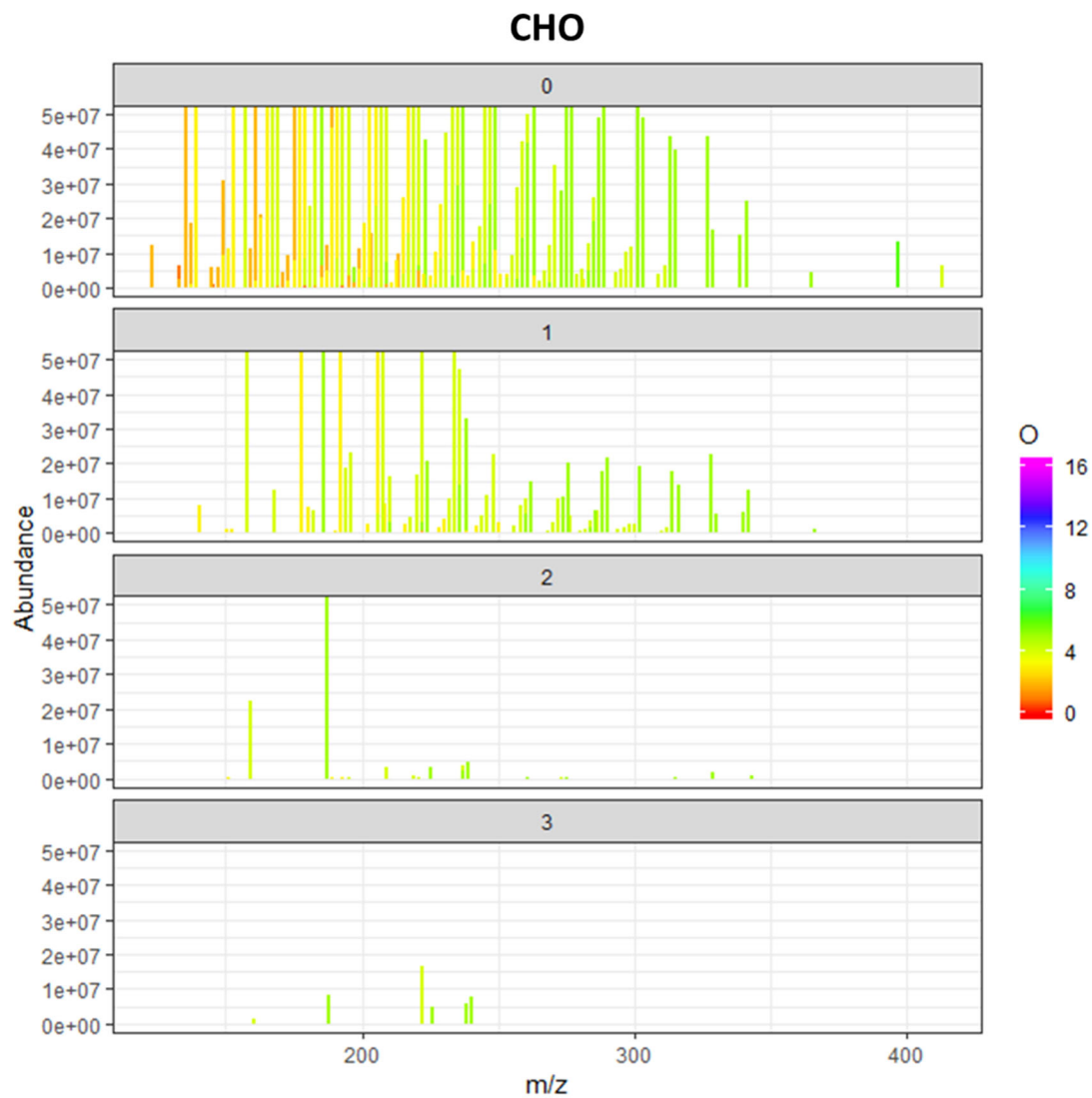


Figure 3.18. Reconstructed mass spectrum of Na^+ adducts after final filtering of MLS MeOH- d_1 data. Spectra are separated by D and elemental group and colors indicate number of O in the formula.

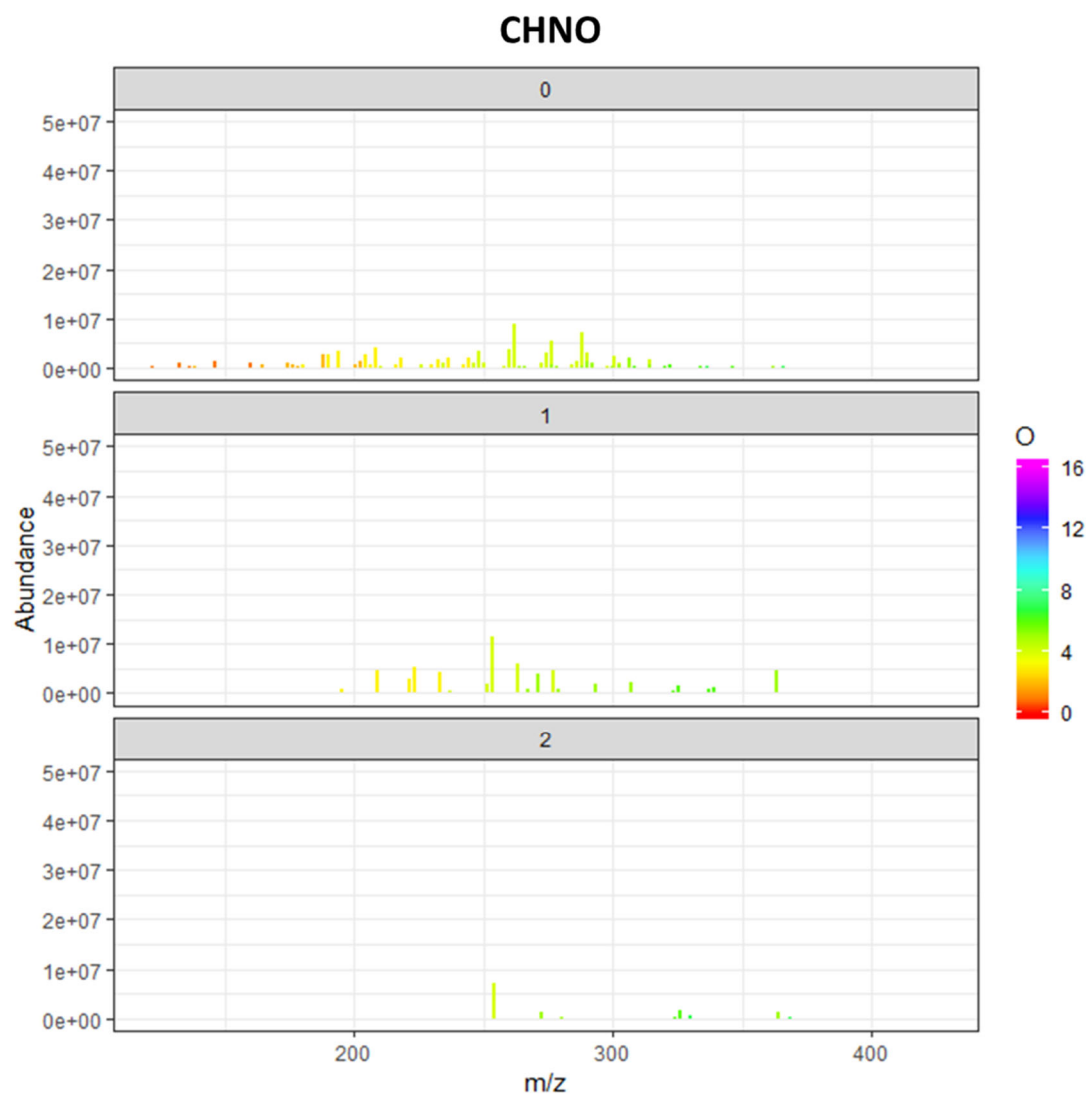


Figure 3.18 (cont.) Reconstructed mass spectrum of Na^+ adducts after final filtering of MLS MeOH- d_1 data. Spectra are separated by D and elemental group and colors indicate number of O in the formula.

The histograms illustrating the O, N, DBE, and C trends of Na^+ adducted MF are given in Figure 3.19(a-d) with the same parameters as those with H^+ adducts. Since there are fewer Na^+ adduct molecular formulas, the distributions of the 2 and 3 exchanges are more scattered. As with H^+ adducts, there are very few D_3 formulas. Despite the few

HDX species available, there is a noticeable shift towards higher O for increasing numbers of D (Figure 3.16a) and there are many molecular formulas with 2 or fewer O atoms that do not participate in exchange. There is also a sharp drop in the frequency of formulas at 6 oxygens. The reason for this is unclear and could be investigated in future works.

Most Na⁺ adducted molecular formulas involved in HDX did not contain N atoms (Figure 3.19b), with about 50% of the non-N formulas undergoing HDX and only about 25% of N₁ Na⁺ adducted MF undergoing exchange. In addition, any MF capable of 3 H/D exchanges did not include N atoms. The DBE distribution of Na⁺ adducts in Figure 3.19(c) is more normal for D₀ and D₁, though distributions for D₂ and D₃ are more scattered due to low number of species. Because of this, the trends in DBE are difficult to interpret. The C trends in Figure 3.19(d) are equally difficult to interpret, although they are slightly right-skewed with a mode around 13-14 carbons.

The molecular formula averages for Na⁺ adducts are detailed in Table 3, similarly as Table 2. Less than half of Na⁺ adducted MF showed any evidence of HDX, and less than 2% showed 2 or more exchanges. Due to the small numbers, there are not enough D₃ formulas to infer the HDX behavior. The DBE tended to increase with increasing D substitution while the average mass increased, with the latter being in contrast to the H⁺ adducts. As with the H⁺ adducts, oxygen and COOH content increased with increasing HDX with the COOH content being calculated identically to that of the H⁺ adducts. The DBE-COOH values are higher for fewer HDX, indicating these species may be rich in

alkenes. This is supported by the AI_{mod} , which indicates most HDX-capable species are olefinic, as are the non-deuterated molecular formulas.

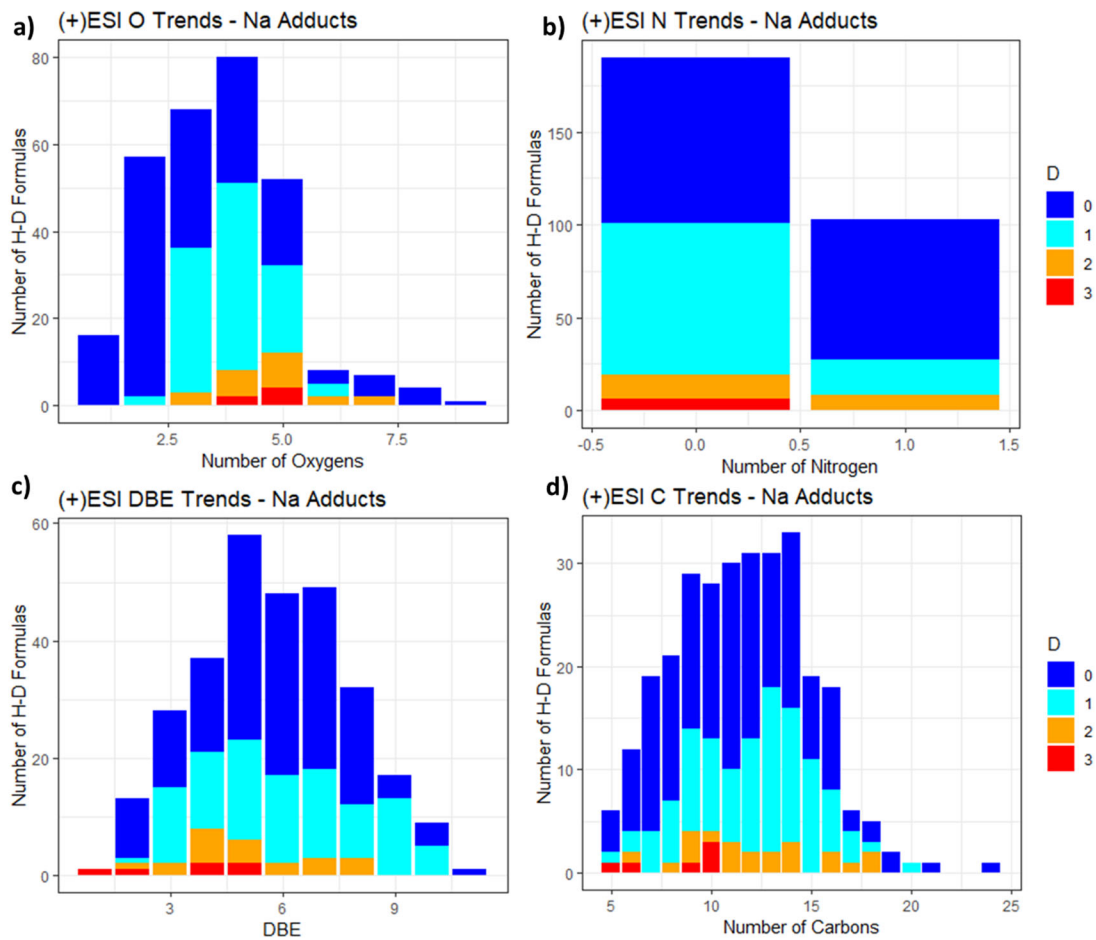


Figure 3.19. Various molecular formula trends with the number of deuterium for all formulas. Colors are based on the number of deuterium. Plots represent number of D per a) oxygen; b) nitrogen; c) DBE; d) carbon.

Table 3.3. Various parameter averages for Na⁺ adducted MF in positive ESI.

D	Formulas	%Formulas	Mass	O	N	O_C
0	293	100.00%	215.74	3.59	0.35	0.32
1	128	43.70%	227.34	4.06	0.21	0.36
2	27	9.20%	234.44	4.7	0.3	0.44
3	6	2.00%	187.41	4.67	0	0.6
D	H_C	DBE	ZarkCOOH	DBE-COOH	Al_{mod}	
0	1.22	5.77	1	4.77	Olefinic	
1	1.23	5.8	1.16	4.64	Olefinic	
2	1.37	4.81	1.4	3.41	Olefinic	
3	1.47	3.5	1.74	1.76	Olefinic	

*An estimate of a sample's overall COOH content, proposed by Zark et al. (2017).

**The difference between DBE and ZarkCOOH.

The van Krevelen plot in Figure 3.20 is set up identically to Figure 3.17, plotting the H/C and O/C ratio of Na adducts against the number of deuterium. Averages for each D are given in filled triangles. Though there are fewer peaks, there is a noticeable trend of increasing H/C and O/C for increasing D substitution.

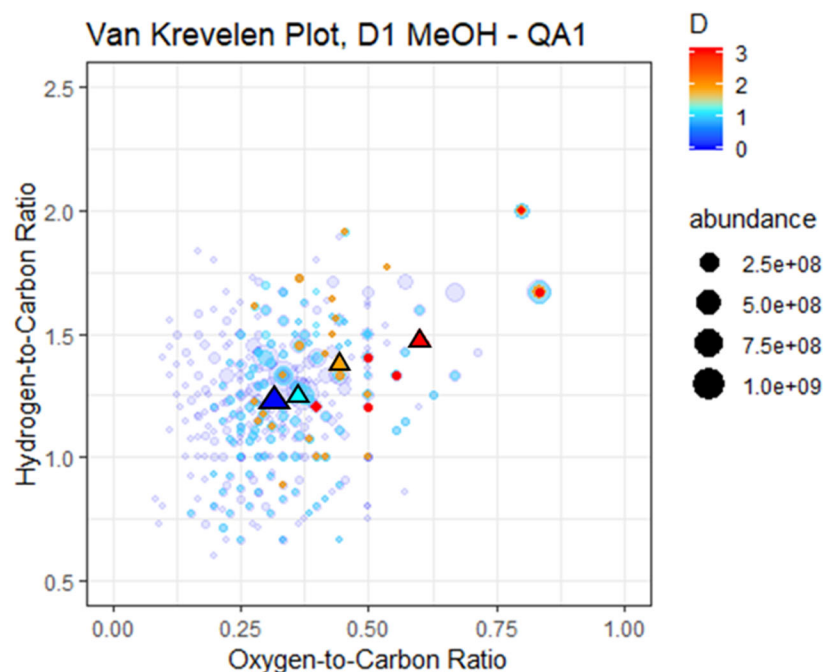


Figure 3.20. Van Krevelen plot for Na^+ adducts of filtered MLS data in positive ESI. Colors refer to the number of D in the formula, and size corresponds to abundance. The colored triangles are the average O/C and H/C for each D as given in Table 3.

3.4 Comparison of Negative and Positive HDX-Capable Species

Detected positive and negative MLS ions show some differences in HDX-capable species. Negative CHO ions had high rates of HDX in comparison to negative CHNO ions, as indicated by the distribution of ions for each number of D post-filtering (Figure 7). This is partially due to the significant presence of $-\text{COOH}$, where there were between 3-6 COOH groups in HDX-capable species (Table 3.1). However, it is noteworthy that due to the ionization mechanism, some acidic species with labile protons, such as organonitrates, may not have been detected. In addition, negative ion N is typically oxidized (Fry et al. 2009; Farmer et al. 2010), and many N species may not have labile protons to perform the exchange.

In negative ESI, 7.5% of the molecular formulas were found with 3 D atoms (Table 3.1). This indicates a reliable number of molecular formulas detected in negative mode for MLS have at least 3 labile protons. However, based on the acid mix experiment, there may be more protons that could display HDX than was observed. For example, PAA showed many instances of HDX, but never had all of its protons exchanged, implying that there is an inherent limit to how much exchange can occur under these conditions. Thus, the species that can undergo 3 exchanges are likely carboxyl-rich, which is supported by the increase in COOH content in Table 3.1. The difference between DBE and COOH decreased, indicating that species with 1 or 2 exchangeable protons may be influenced by alkenes. Specifically, enol groups could be responsible for this trend, where COOH content is not necessary to perform the exchange. An example of this process in MeOH-d₁ is given in Figure 3.21a. Tautomerization could occur with carbonyl compounds (or alcohols) to move unsaturation between keto and enol forms. Thus, instead of having labile protons on a heteroatom, the labile protons may be on resonance-stabilized carbons. This could be an explanation for HDX occurring with olefinic species, though it is difficult to validate this without more structural information.

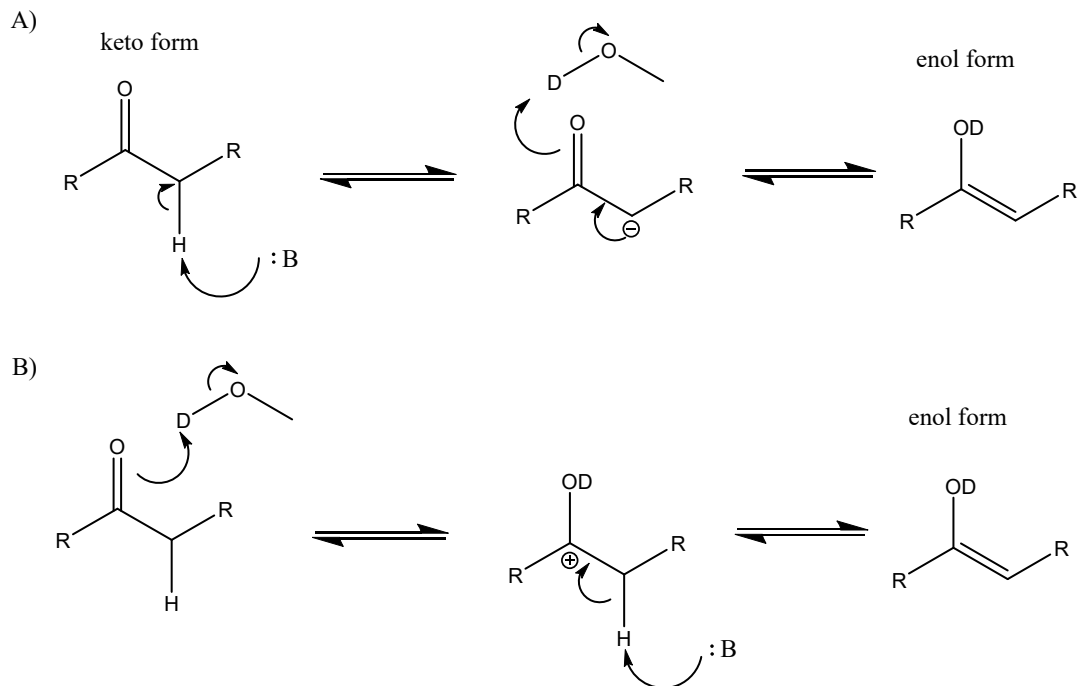


Figure 3.21. Mechanism for keto-enol tautomerization under A) basic conditions; B) acidic conditions.

Several parameters may be potential predictors for HDX ability. Along with the increase in O content, the average mass increases with increasing D substitution for negative ESI. This is reasonable since the greater the number of O atoms, the greater the mass of the compound and the more likely there are to be labile protons. In addition, the H/C also increases, implying compounds that are more saturated have an enhanced HDX ability. This is reiterated in the van Krevelen plot in Figure 3.9, where the average O/C and H/C values increase linearly with an increase in H/D exchangeability. Clearly, both H/C and O/C are important factors in determining a compound's propensity to undergo HDX, and could be used as predictors for the functionality of a particular formula.

Due to the general decrease in DBE, it is possible that enol-type compounds are important to lower amounts of HDX while carboxylic groups are important at higher amounts of HDX. However, there are several D_3 formulas that have DBE values of less than 3 (Figure 3.8b), and as such, may not be fully explained by carboxyl groups. In addition, there are a few instances of HDX where there are more exchanges than carboxyl groups could account for. For example, there are some D_1 formulas that are only associated with 1 O atom. Since carboxyl groups contain 2 O atoms, the HDX in this case cannot be from a carboxyl. If the O were part of an enol, however, the tautomerization could explain the larger number of HDX due to the relative lability of the α -C-H bonds.

The positive ion mode of ESI has the added complication of more molecular formulas with different types of adducts, typically H^+ and Na^+ . Comparison of the two indicates that the two types of molecular formulas behave differently in terms of HDX. Na^+ adducted compositions in positive ESI are primarily CHO compounds (Figure 3.18). However, the CHO compounds observed here have lower numbers of O and have a much lower estimated COOH in comparison to negative mode. The overall slope of the O/C vs. H/C line in the van Krevelen plot containing Na^+ adducted species (Figure 3.20) is very similar, however, it is shifted to lower O/C ratios. This implies that while the number of O is important for the HDX behavior in positive ions, a lower number of O atoms are needed to perform multiple exchanges. In this case, the Na^+ adduct may be detecting the non-acidic oxygen atoms, such as ether and ketone functional groups. These ketones could potentially be converted to $-OH$ through acidic HDX of an enol, as shown in Figure 3.21b.

Based on the reconstructed mass spectra of formulas in Figure 3.15, O does not play a major role in HDX for H^+ adducted species. Due to the low number of CHO compounds with H^+ adducts, it is to be expected that the O, O/C, and COOH content are low. It is possible that compounds are not undergoing HDX, but are instead being protonated with the labile D from MeOH- d_1 . Kostyukevich *et al.* investigated this case with their work and found that all species were ionized by a deuterium rather than a proton in their deuterium-rich atmosphere. However, the use of methanol instead of water may have prevented this from occurring so readily, as not all detected ions were deuterated.

Instead, N is the primary contributor to HDX in positive mode, which reiterates the preferentiality of ionizing more basic compounds like amines with positive ESI. In addition, this mode highlights the high N content of MLS, with at least 50% of formulas being detected in the positive ion mode containing N (Table 3.2). To our knowledge, this detailed of an investigation of HDX on N has not been explored previously. The most relevant work is by Kostyukevich *et al.* (2013) who did some work with S-containing compounds, but did not emphasize N. In this study, we found the number of CHNO compounds in positive ion mode nearly equivalent to the number of CHO compounds detected in negative ESI.

The number of HDX molecular formulas drops more drastically for H^+ adducts for increasing D, with less than 14% of formulas able to undergo 2 exchanges. This could imply that HDX-capable H^+ adducted species are likely those that contain only one exchangeable H such as secondary amines, amides, or aromatic N like pyridine or

imidazole. Less than 1% of formulas were capable of 3 exchanges, and due to the low number, indicating that the detected species are less likely to have multiple labile protons.

The O/C ratio increases for increasing HDX ability (Figure 3.17), though there is little trend in the H/C ratio. However, the DBE-COOH value decreases. The sample's COOH content increases and its olefinic nature remains approximately constant for varying degrees of HDX as reported in Table 3.2. This is supported by the tautomerization mechanism in Figure 3.21. In addition, this mechanism occurs under both acidic and basic conditions, implying it could be facilitated with either electrospray polarity. However, more analysis would be required to confirm this.

4 Characterization of Ammonium Artifacts Formed Using a 2-Step Solid-Phase Extraction Method with Liquid Smoke

The following section provides a summary of the various samples, ionization modes, and analyses performed to investigate the extent of the $^{15}\text{NH}_4^+$ artifact. The results from negative and positive ESI data will be discussed separately using similar analyses. First, the SPE-prepared samples are discussed, where the general appearance of the samples after the initial molecular formula assignment will be considered to understand their general composition. The assigned molecular formulas were then filtered and their ambiguity was assessed to simplify the data set. Next, several potential reactive pathways are proposed for artifact formation. The precursors for these reactions were matched to further validate assigned ^{15}N molecular formulas and determine the most likely reaction pathways. Then, the non-SPE samples were analyzed to investigate the effects of pH on artifact formation.

4.1 Negative ESI Analysis

4.1.1 Initial Molecular Formula Assignment

Negative ion mass spectra for the first elution (L1) from the SPE cartridge using 90:10 MeOH:H₂O of MLS and HLS are given in Figure 4.1. These spectra highlight the similarities between the two samples. The intensity scale is adjusted to an intensity of 25,000 to visualize the ion distributions of the lower intensity ions. CH and CHN formulas were also assigned, but were omitted due to their low frequency (Smith et al.

1990). Visually, the two samples are very similar in their spectral complexity with high intensity ions below m/z 400 in patterns comparable to biomass burning samples (Mazzoleni et al. 2012). The CHO ions in this range have DBE values up to about 15. Although lower in intensity, each sample also has similar ion distribution patterns of CHNO compounds. More abundant ions were detected in the m/z 250-350 range, and a set of ions with a Gaussian-like distribution were detected from m/z 350 onward. These ions tended to have lower DBE values. The taller, mid-DBE CHNO ions in HLS are well-known contaminants within the MeOH solvent and were omitted from analysis. In total, about 70% of the detected ions in both MLS and HLS L1 were CHO. In addition, approximately 12% more ions were detected in MLS than HLS (1968 vs. 1730, Table 4.1), a trend observed in all of the different sample preparations. Due to the larger number of ions and great similarity between the samples, the following results and discussion will be focused on MLS and HLS can be assumed to be similar except where noted.

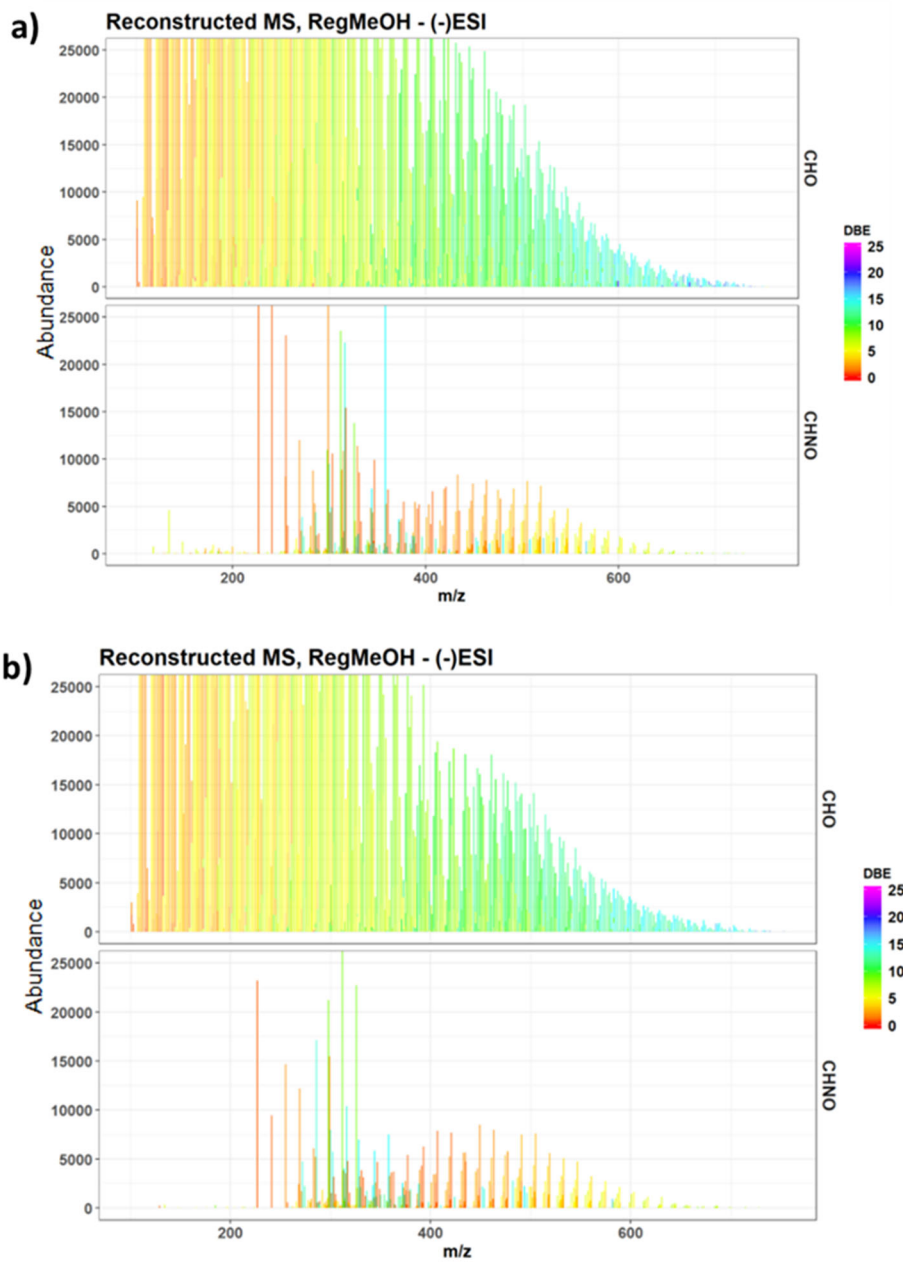


Figure 4.1. Reconstructed mass spectra for initial assignment of first elution (L1) liquid smoke samples. a) Mesquite liquid smoke; b) Hickory liquid smoke. Mass spectra are separated by group and colored by DBE.

Initial MLS mass spectra for the second elution (L2) from the SPE cartridge with 0.3% NH₄OH in MeOH are given in Figure 4.2. The mass spectra are separated by elemental group in addition to the type of NH₄OH used. The acidified and non-acidified samples are also provided for comparison. The mass spectra for NH₄OH and ¹⁵NH₄OH samples looked similar with slight differences in the overall ion intensity. Due to the more dilute nature of L2 relative to L1, the methanol contaminant peaks are quite strong for these samples. However, the CHNO ions otherwise have a wide spread across the mass spectrum with the Gaussian-like distribution at a maximum just below m/z 500. The major difference between L1 (Figure 4.1) and L2 is the overall increase in DBE for L2, especially for CHNO compounds.

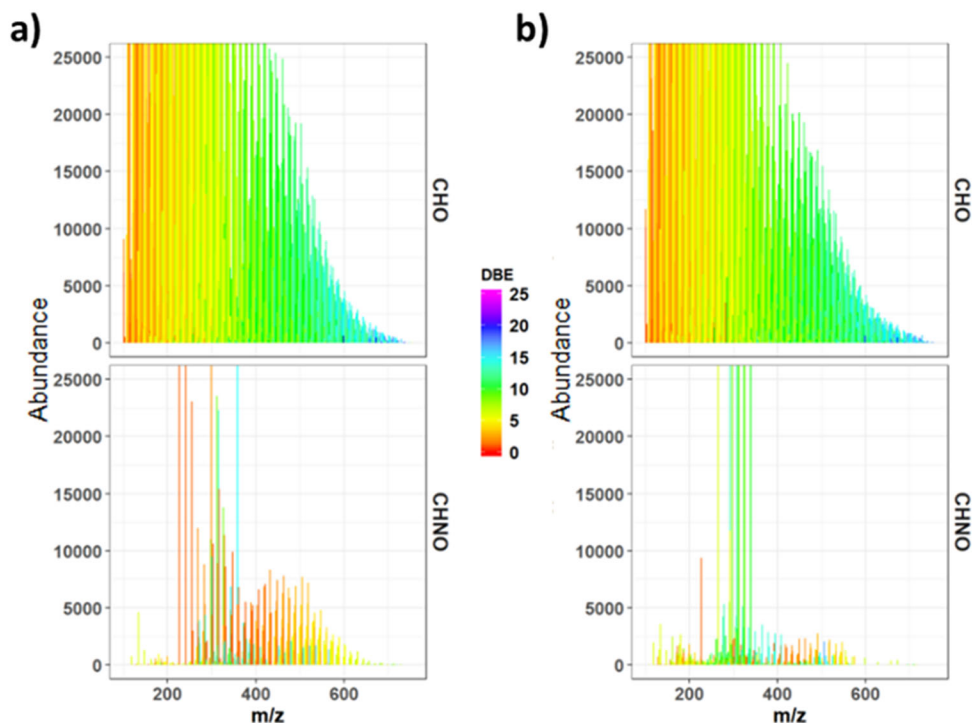


Figure 4.2. Reconstructed mass spectra for MLS second elution (L2) samples. All mass spectra are divided by group and colors indicate DBE. a) Regular NH_4OH without acid; b) $^{15}\text{NH}_4\text{OH}$ without acid.

4.1.2 Ambiguity and Quality Assurance

Before filtering and analyzing the data further, it was necessary to ensure there were no ambiguous assignments in the regular NH_4OH samples, since the $^{15}\text{NH}_4\text{OH}$ data will ultimately be compared to it in order to remove unlikely formulas. Ambiguous assignments were limited to fewer than 50 in all NH_4OH samples and a majority of them involved a mass difference between C_{12} and $\text{H}_4\text{O}_7\text{N}_2$ (mass difference of 0.0019 Da). In these cases, the molecular formula with fewer N was selected based on several factors: 1) fewer heteroatoms tend to give more chemically feasible formulas (Koch et al. 2007); 2)

CHO compounds are more likely to be detected than CHNO compounds in negative ESI due to poor N acidity; 3) the DBE values for formulas with fewer N are higher, and more representative of the overall sample; 4) the DBE-O values are closer to 7 or 8 as opposed to -12 or -13, which are more chemically reasonable (Herzprung et al. 2014); 5) the absolute error for the CHO assignments was lower than the CHNO assignments. After selecting the appropriate ambiguous assignments, the resulting list of formulas was considered the final list for the control samples.

The ^{15}N samples were then filtered for only molecular formulas present in the control sample, as described in Section 2.3.2. This filter greatly reduced the number of $^{15}\text{NH}_4\text{OH}$ ions to less than the corresponding control data. There were no remaining ambiguous molecular formulas. The result of this filter is given in the van Krevelen plot in Figure 4.3. Figure 4.3a represents MLS L2 without acid, and Figure 4.3b represents the same with 0.1% acid. A majority of the CHNO molecular formulas with a high O/C and H/C were removed indicating the filter was effective. In addition, there is a noticeable similarity between samples with and without acid.

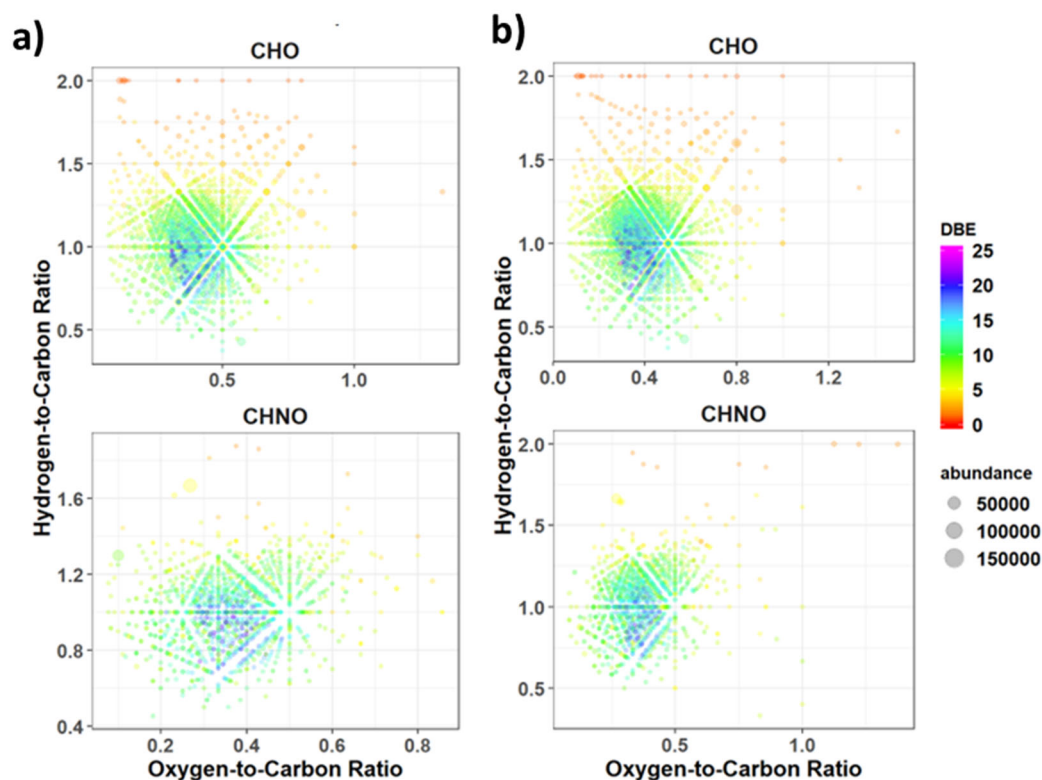


Figure 4.3. van Krevelen plots of MLS L2 with $^{15}\text{NH}_4\text{OH}$ molecular formulas after filtering. Plots are divided by group and colors indicate DBE. a) Sample without added acid; b) Sample with added acid.

4.1.3 Composition of ^{15}N Molecular Formulas

The reconstructed mass spectra of the filtered results are given in Figure 4.4. It is clear that the species with 2 or 3 ^{15}N atoms are less reliable as they are low in abundance and scattered. Molecular formulas containing only one ^{15}N have a bimodal distribution pattern, similar to what is seen in the control samples. This distribution is potentially due to oligomerization of lower molecular weight species, (Sun et al. 2010; Yasmeen et al. 2010) and is clearly present in the isotopically-labeled samples. Figure 4.5 also indicates

the general absence of formulas in negative ESI containing 2 or 3 ^{15}N . There is also a faint trend in the $^{15}\text{N}_2$ formulas with error inversely proportional to the formula mass. This indicates that the artifact does not significantly occur multiple times for the species observed as negative ions. As such, $^{15}\text{N}_2$ and $^{15}\text{N}_3$ molecular formulas are not further considered.

Although the ^{15}N formulas were matched to a molecular formula in the ^{14}N samples, this does not necessarily mean the formulas are completely reasonable; they may still have unusual trends in terms of their error outside of the systematic instrument error. If the errors for the ^{15}N molecular formulas do correspond to the error trends of the other species, then the ^{15}N molecular formulas are more likely to be reasonable. The absolute error vs. mass plot is given in Figure 4.5 for MLS L2 with formic acid, sectioned by the number of ^{15}N in the formula and colored by the absolute error to highlight distribution patterns. This sample was selected as an example because all other MLS and HLS samples with and without acid were similar. There are a significant number of molecular formulas in this sample with one ^{15}N and the spread of the masses in the error plot is relatively random, closely resembling the trends of the other ions. It can then be concluded that these assignments are reasonable.

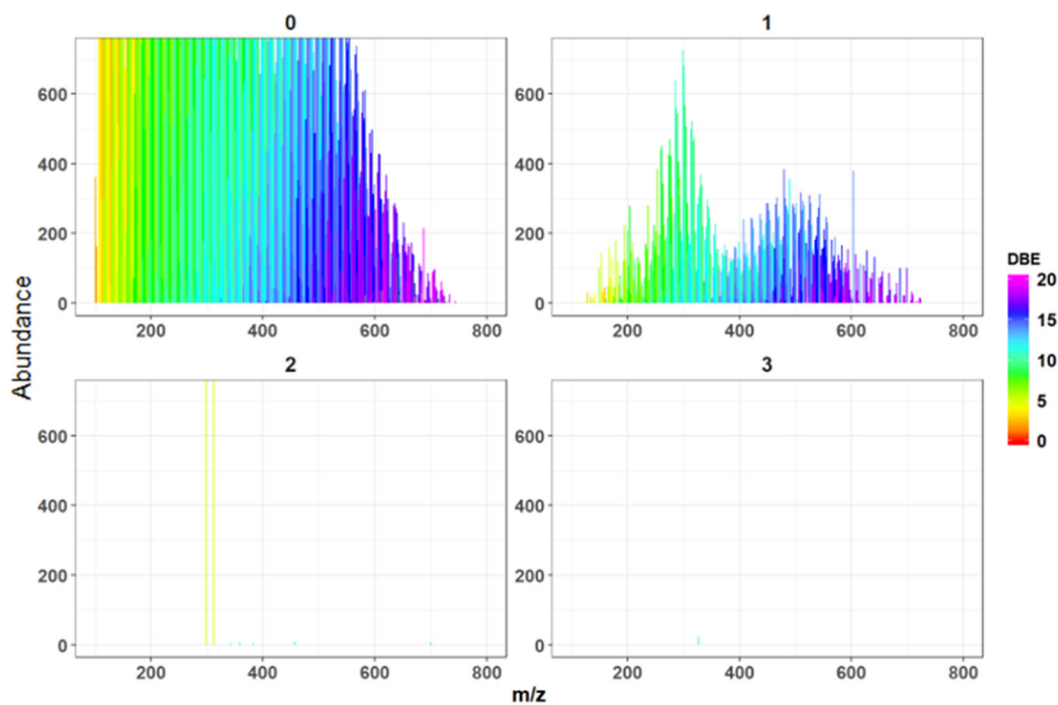


Figure 4.4. Reconstructed mass spectra of MLS L2 in $^{15}\text{NH}_4\text{OH}$ with acid molecular formulas. Colors indicate DBE and divisions are by formula based on the number of ^{15}N atoms assigned.

The breakdown of CHNO molecular formulas for each $^{15}\text{NH}_4\text{OH}$ sample is summarized in Table 4.1. This table is organized by the sample type and contains the total number of molecular formulas assigned after filtering against the control sample. Each sample is divided by the N-isotope; ^{14}N indicates the assigned CHNO molecular formulas that contain only ^{14}N atoms, and ^{15}N indicates the assigned CHNO molecular formulas containing a ^{15}N atom. In this case, the ^{14}N molecular formulas can be considered representative of the “natural” amount of CHNO in the sample, while ^{15}N molecular formulas are indicators of an artifact occurring. The CHNO MF and % Total MF represents the number of either CH^{14}NO or CH^{15}NO molecular formulas and the

percentage of these formulas compared to the total number of molecular formulas assigned. In addition, the total ion current (TIC) and average abundance are tabulated. The average abundance is calculated as the average measured abundance of all CH^{14}NO or CH^{15}NO ions and the TIC is the sum of the same values. These values are general measures of how significant these particular ions are within the mass spectrum.

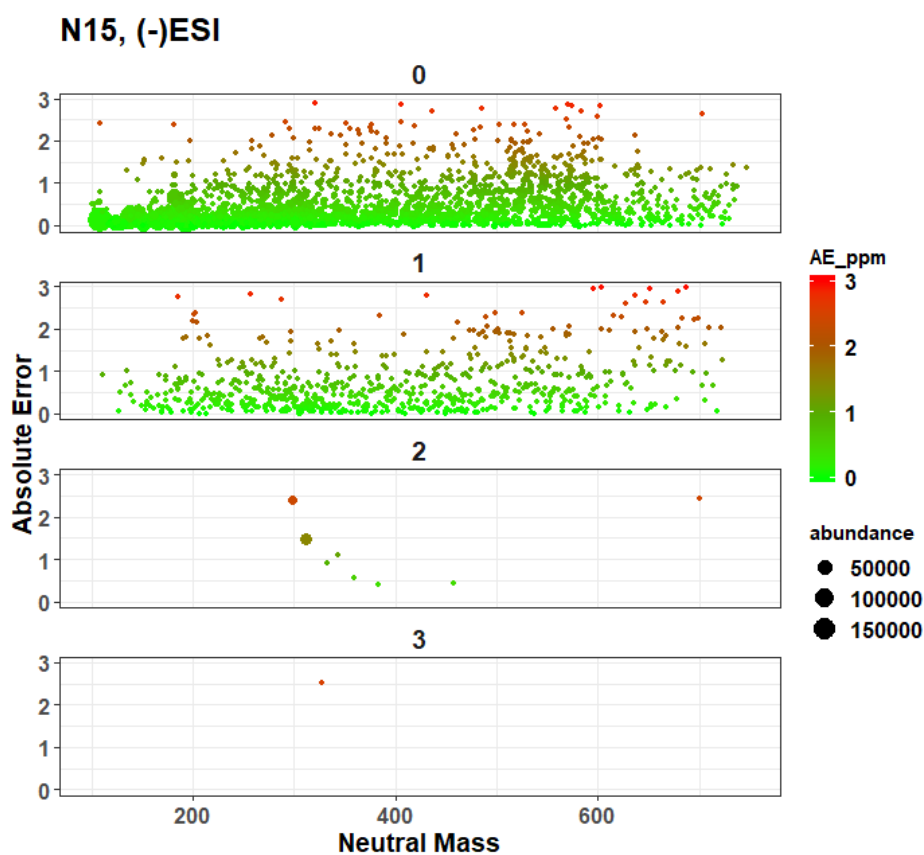


Figure 4.5. Absolute error vs. theoretical mass of MLS L2 in $^{15}\text{NH}_4\text{OH}$ with acid molecular formulas. Colors indicate DBE and plots are divided by formula based on the number of ^{15}N atoms assigned.

For all samples, 35-45% of assigned molecular formulas contained one ^{15}N . This value is significant, as regular ^{14}N accounts for 13-23% of the assigned molecular formulas in all cases. This indicates that a majority of the reasonable CHNO compounds detected for L2 result from an artifact between the sample analytes and the NH_4OH during the SPE process. Additionally, the total and average abundance of the ^{15}N CHNO molecular formulas are between 2-10 times greater than those with ^{14}N . Thus, the NH_4OH artifact is not only significant in terms of number, but also in terms of ion abundance. Although these total abundance values are orders of magnitude less than those of the CHO molecular formulas, these values depict an overall increase in the CHNO content within these samples. This is consistent with the large number of CHNO compounds observed in the regular NH_4OH -treated samples.

Table 4.1. Number of CHNO molecular formulas of different N isotopes for all samples.

Sample Num. MF	N Isotope*	CHNO MF	%Total MF**	Total Ion Abundance	Avg. Ion Abundance
MLS L2, 0.1%FA 1859 MF	¹⁴ N	404	21.7%	18030	44.6
	¹⁵ N	666	35.8%	72262	109
MLS L2, 0%FA 1699 MF	¹⁴ N	401	23.6%	28620	71.4
	¹⁵ N	762	44.8%	338525	444
HLS L2, 0.1%FA 1753 MF	¹⁴ N	220	13.4%	81776	372
	¹⁵ N	615	37.4%	1451887	2361
HLS L2, 0%FA 1643 MF	¹⁴ N	274	15.6%	27913	102
	¹⁵ N	775	44.2%	167187	216

*¹⁴N represents CHNO MF with only ¹⁴N; ¹⁵N contains MF with at least one ¹⁵N.

**Calculated as a ratio of CHNO MF to Num. MF in the first column.

Several molecular trends for the isotopically-labeled MLS L2 with acid sample are summarized in the histograms in Figure 4.6. In all figures, the vertical axis represents the number of formulas with a particular variable such as oxygen number. The colors indicate the number of ¹⁵N atoms assigned in that molecular formula with blue representing formulas without ¹⁵N and cyan representing molecular formulas with one ¹⁵N. In general, the C, O, and DBE trends for ¹⁵N formulas mirror those of the full data set, indicating it is difficult to attach the artifact formation to a particular molecular composition. The log(abundance) trends histogram depicts a mirroring effect as well; the ¹⁵N molecular formulas have a generally lower intensity than non-labeled ones, but they have a very similar distribution. As a result, more steps are required to explain the nature of this artifact.

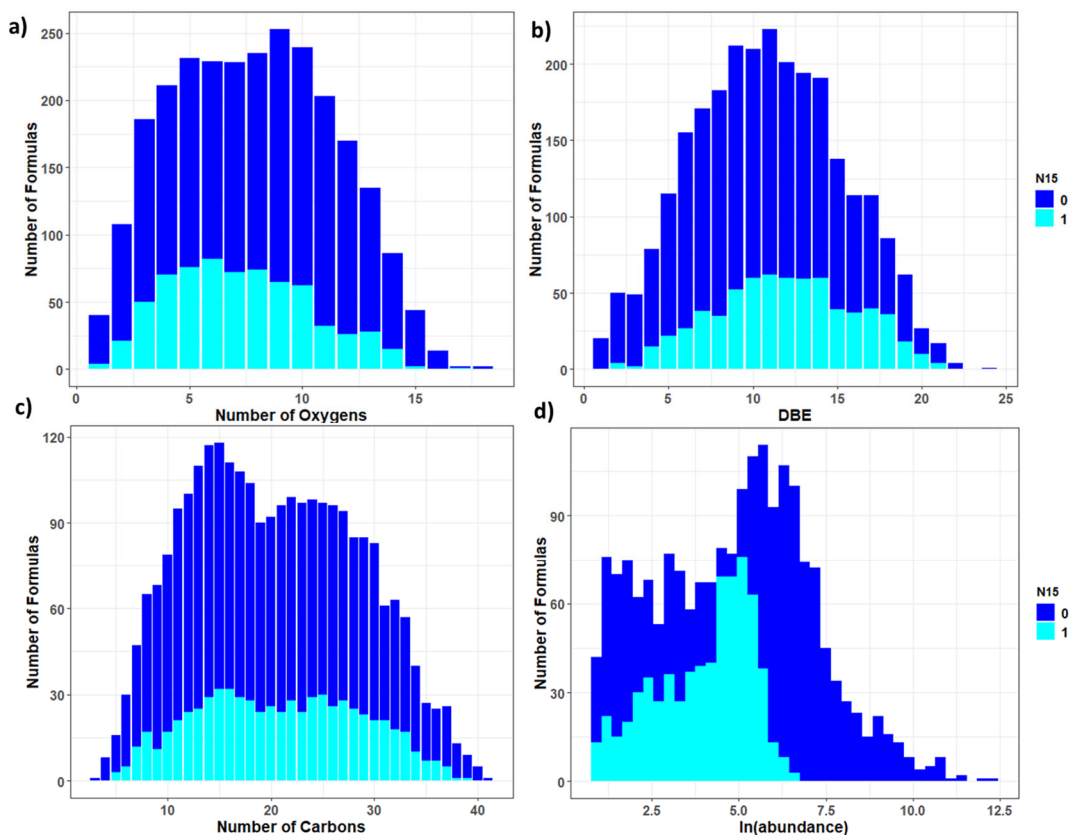


Figure 4.6. Molecular trends for CHNO ions in MLS with acid. Colors in all plots indicate molecular formulas containing 0 or 1 ¹⁵N. All plots measure number of molecular formulas vs. a) Number of O; b) DBE; c) Number of C; d) ln(abundance).

4.1.4 Potential Adduct and Reactive Artifact Pathways

Given the evidence for artifact formation in liquid smoke L2 samples, analysis to determine the nature of this artifact is needed. Ideally, the reaction pathways and precursor species can be determined. Although ammonia is well-known to be a versatile reactant, this preliminary investigation will focus on its interactions with carbonyl compounds. This is because CHO molecular formulas and carbonyl functional groups are important aspects in most biomass burning aerosol. (Baugh et al. 1987; Zhang and Smith 1999)

Typically, the reactions of carbonyls require acidic conditions to be favorable with a moderate nucleophile such as NH_3 . However, since some of the ^{15}N molecular formulas are presumed to be reactive artifacts, the electrospray polarity does not necessarily restrict their formation. Reactions requiring a proton-donating catalyst may still occur prior to the electrospray process, especially since a small amount of water is present in the samples.

In negative ESI, three different artifact forming reaction pathways were studied, all involving NH_3 interacting with a carbonyl compound. These general reactions are summarized in Scheme 1. In Scheme 1a, the oxygen of a ketone or aldehyde is replaced by NH to form an imine (Schiff base), via nucleophilic addition of NH_3 to the carbonyl carbon and eventually transferring protons to the oxygen to eliminate H_2O . Typically, this reaction favors mildly acidic conditions (Reusch 1999). This reaction is referred to as “Imine” for the remainder of this chapter. The reaction in Figure 4.7b is a less favored conversion of a carboxyl to an amide. Due to competing acid-base interactions between NH_3 and the acid, this reaction requires high temperatures ($>100^\circ\text{C}$) in order to proceed forward, which may be reached within the electrospray, but is unlikely to occur in solution. In addition, it is difficult to differentiate between this reaction and the reaction in Figure 4.7a, since the mass difference from precursor to product is identical in both cases (Loss of OH and addition of NH_2). Figure 4.7c is theoretically the most likely reaction to occur, because it does not require an acid catalyst. This reaction pathway converts esters to amides via nucleophilic substitution, releasing a stable alcohol as a leaving group. A methyl ester is used as an example below, but this reaction can occur

with a variety of esters. This reaction route will be labeled as “Ester” for the remainder of this chapter.

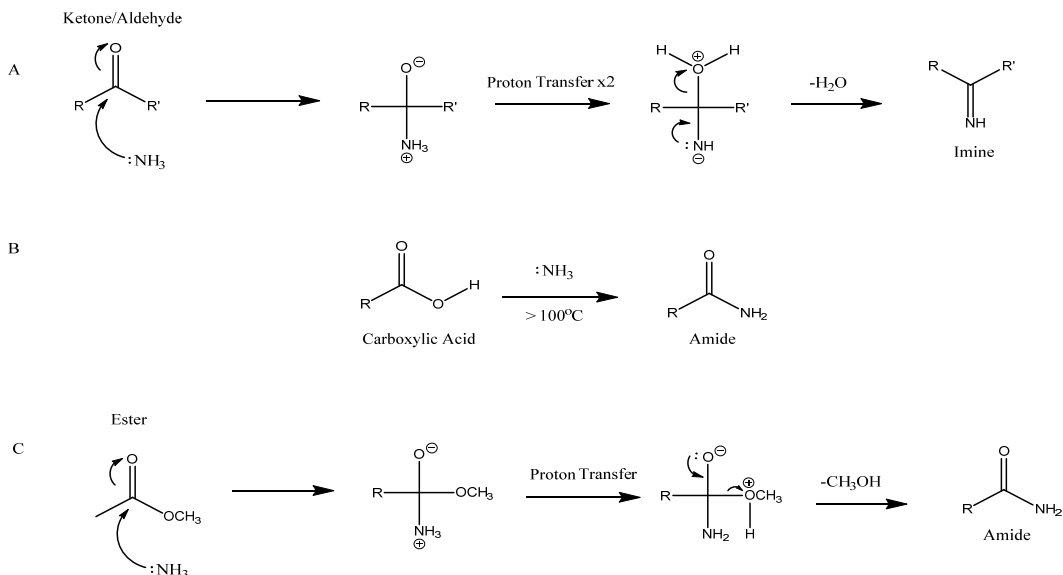


Figure 4.7. Potential reactive pathways for producing ammonia artifacts in liquid smoke.

To determine if these particular reactions occurred, the ^{15}N formulas were screened for the presence of one of the viable precursors within the MLS L2 0.1%FA sample as an example, as results for other samples were similar. For each reaction in Figure 4.7, the theoretical exact mass difference between the precursor and product was calculated and used to locate the potential precursors in the sample. For the $^{15}\text{NH}_4\text{OH}$ samples, the mass differences were based on the addition of ^{15}N , while ^{14}N was used for the control samples. The mass differences for the reactions involving both $^{14}\text{NH}_3$ and $^{15}\text{NH}_3$ are given in Table 4.2. These mass differences were calculated as product mass minus the precursor mass. As stated previously, the mass differences for the imine

formation and the carboxylic acid hydrolysis are identical, so they cannot be differentiated in this manner. As such, due to the high energy requirements for the acid reaction, this analysis will be focused on the more favorable imine formation. For the ester to amide conversion, only the mass differences for methyl esters (OCH₃) were considered. Although not included here, this procedure could easily be extended to other esters.

Table 4.2. Mass differences between the organic products and reactants for reactions in Figure 4.7 used to locate potential precursors for ¹⁵N artifacts.

Reaction Pathway	Elements Added	Elements Removed	N14 Mass Difference	N15 Mass Difference
Imine Formation	NH ₃	H ₂ O	-0.98402	0.01302
Acid to Amide	NH ₃	H ₂ O	-0.98402	0.01302
Methyl Ester to Amide	NH ₃	CH ₃ OH	-14.99967	-14.00263

Once the ¹⁵N samples were screened for precursors in each reaction pathway, an additional level of QA was performed to keep only molecular formulas most likely to occur from that pathway. The screening was performed on the control samples using ¹⁴N mass differences. Similar to the QA filter applied to the ¹⁵N data, the exact precursor/product reaction pairs needed to be present in the control samples as well. Following this filter, of the original 666 ¹⁵N formulas, 586 could be traced to imine formation and 586 could be traced to a methyl ester hydrolysis. Surprisingly, these two

numbers were identical. Note that there is significant overlap between the two pathways, indicating some molecular formulas could have resulted from one or the other, or potentially both.

In addition, the relative abundances of the product and the precursor were compared to determine the average extent of the equilibrium for artifact formation. Since there is such a breadth of molecular formulas and the exact conditions during the electrospray process are not well understood, there were no preconceived expectations. The boxplot distribution of the ratios of artifact abundance to precursor abundance are given in Figure 4.8. The red dot represents the mean abundance ratio for the sample. The left figure represents the ratio distribution for the imine formation pathway for the $^{14}\text{NH}_4\text{OH}$ sample, and the right represents the same for the $^{15}\text{NH}_4\text{OH}$ sample. The vertical axis is scaled by the natural log of the abundance ratio, so positive numbers indicate a more abundant artifact ion while negative numbers indicate a more abundant precursor ion.

Both the ^{15}N and ^{14}N samples have nearly identical distributions, again confirming that the NH_3 artifact is replicable between the ^{14}N and ^{15}N samples. In addition, the distributions are each centered around a ratio of 2.5 with more than 75% of the data greater than 0, indicating that the reaction artifact for this proposed pathway is typically more abundant than the precursor. This implies that the equilibrium for this process favors the N-containing products, and that the reaction artifact significantly decreases the abundance of the present CHO precursor when it should be the dominant species. These trends also hold for the ester hydrolysis data.

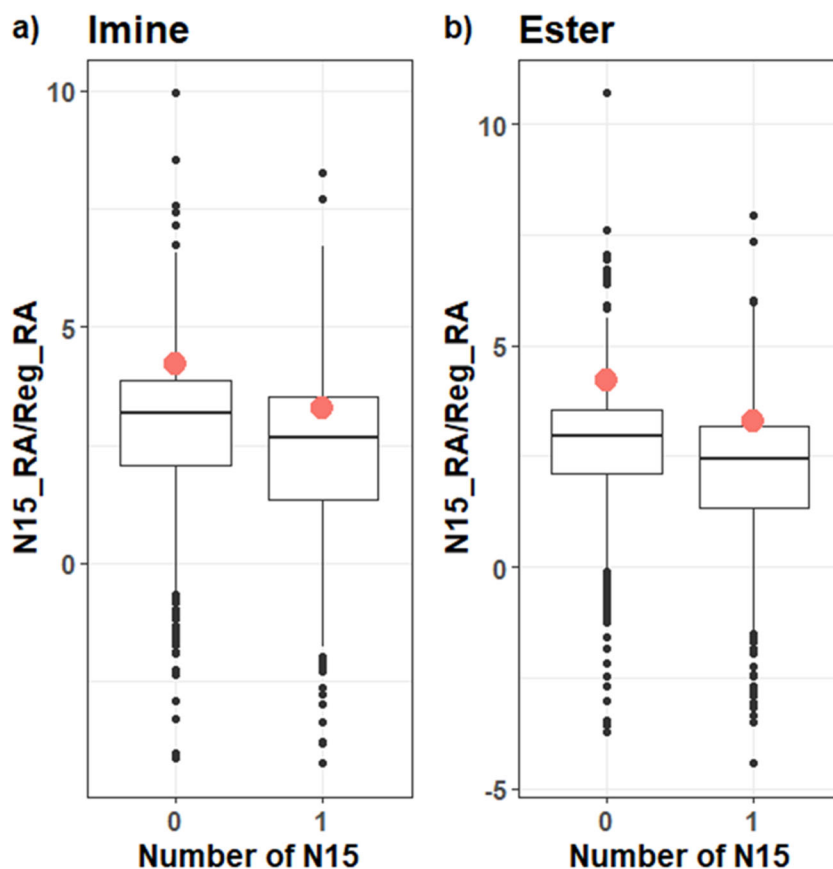


Figure 4.8. Boxplots of the artifact-to-potential precursor abundance ratios in acidified MLS samples for a) imine formation; and b) ester hydrolysis. 0 and 1 N15 refer to the control and labeled NH_4OH , respectively.

To go one step further, it was desirable to assign potential artifact formulas to one particular reaction pathway when both pathways were possible to differentiate molecular formulas that may be more likely one or the other. Though there was no definitive method for doing so in such a complex mixture, artifact molecular formulas could be classified by their likelihood based on a comparison of their relative abundances. For molecular formulas where a potential artifact could be the result of both imine formation and ester hydrolysis in the ^{15}N sample, the ratio of product/precursor abundance for each reaction was compared via a second ratio to the same reaction and product/precursor pair

in the regular sample. The reaction with this second ratio closer to 1 was then selected as the “assigned” reactive pathway for this artifact, as this pathway was more representative of abundance ratios in the regular NH₄OH sample. Then, the pathway was assigned with a qualifier based on how different the abundance ratios of the two reaction pathways were. The meaning of these qualifiers are summarized in Table 4.3.

Table 4.3. Definitions of qualifiers used in comparing likelihoods of each reaction.

Qualifier	Description*
<i>Unambiguous</i>	Only one of the reactive pathways leads to this artifact.
<i>Good</i>	The percent difference in the abundance ratios in comparison to the control is > 100%
<i>Likely</i>	The percent difference in the abundance ratios in comparison to the control is 50-100%
<i>Possible</i>	The percent difference in the abundance ratios in comparison to the control is 10-50%
<i>Inconclusive</i>	The percent difference in the abundance ratios in comparison to the control is < 10%

*Percent differences are a measure of how precursor-artifact abundance ratios compare between ¹⁵NH₄OH and ¹⁴NH₄OH for each reaction pathway.

It is clear that this step is a large generalization that makes several assumptions. Although only the ratio closest to 1 is selected as the final product, it is entirely possible that both reactions do, in fact, occur in the same sample to yield the same CHNO molecular formula. In addition, this assumes that 100% of the ¹⁵N peak is caused by a reactive artifact and 100% of the corresponding peak in the regular sample is due to the artifact as well. Although the former may be a reasonable assumption, the latter

assumption may not be, as there may be some naturally occurring amount of a particular CHNO compound that did not come from a reaction during SPE. The purpose of this investigation was to find any differences in the molecular parameters between the two reactive pathways, given molecular formulas that were *likely* to be associated with one pathway or the other. Because liquid smoke is such a complex mixture, it is still quite difficult to determine the true source of these artifacts. Though there are several precursor peaks that lead to valid formation pathways, confirming these pathways is nearly impossible for such a sample.

The results from this tagging and analysis are depicted in Table 4.4, where the averages of several molecular parameters are calculated for the potential artifact precursors in both MLS and HLS. The table is separated based on the reaction pathway as well as the likelihood of this reaction being dominant over the other pathway. Only the top three categories (*Unambiguous, Good, and Likely*) are included to contrast the two reaction pathways as much as possible. Despite this, there are great similarities between the two reactions based on these loose definitions. Across all likely reaction pathways, there is little difference in the molecular characteristics, even though there is a difference in the number of molecular formulas capable of undergoing each pathway, especially in the case of HLS. Interestingly, there are more Imine molecular formulas in MLS, but significantly more Ester formulas in HLS. The most notable differences between pathways are in HLS, where Ester precursors have a slightly higher average mass and number of oxygen. In MLS, this relationship is reversed and the Imines have more

precursors with a higher average mass and higher number of O. Clearly, this indicates subtle molecular differences between HLS and MLS.

Unambiguous molecular formulas for each pathway also indicate differences between the reactive species. Since there are so few *Good* and *Likely* formulas, these were not considered. Although the overall trends in theoretical mass and O number were different for MLS and HLS, both samples show the same trends for *Unambiguous* formulas, where the molecular mass and the number of oxygens are greater for the Imine reactions. In addition the DBE is slightly higher, indicating more unsaturation in the Imine precursors. This appears counterintuitive since Ester precursors naturally have two O atoms per functional group, while Imine precursors (ketones and aldehydes) have only one. This could either indicate that this method of separating reactions provides a weak representation of the molecular formulas potentially involved or that there are other factors that go into determining the most likely reactive pathway that have not been considered.

Table 4.4. Averages and counts of molecular formula parameters in MLS and HLS L2 with acid for precursors of each reaction pathway and likelihood.

MLS - L2 with Acid			Averages					
Rxn	Likelihood	Formulas	Theoretical Mass	O	N	O/C	H/C	DBE
Ester	All	293	381	7.60	0.01	0.39	1.00	11.1
	Unambiguous	71	392	7.76	0.04	0.39	0.98	11.6
	Good	10	489	10.80	0.00	0.44	0.83	14.9
	Likely	3	312	7.33	0.00	0.49	0.74	10.7
Imine	All	325	406	8.33	0.01	0.41	0.97	11.7
	Unambiguous	77	403	8.30	0.04	0.41	0.92	12.2
	Good	11	428	9.27	0.09	0.45	1.05	11.5
	Likely	10	539	11.80	0.00	0.44	0.93	14.7
HLS - L2 with Acid			Averages					
Rxn	Likelihood	Formulas	Theoretical Mass	O	N	O/C	H/C	DBE
Ester	All	445	411	8.24	0.01	0.39	1.03	11.3
	Unambiguous	96	442	8.93	0.02	0.39	1.00	12.4
	Good	26	467	10.30	0.04	0.44	0.89	13.6
	Likely	14	368	8.21	0.00	0.44	0.98	10.6
Imine	All	248	397	7.90	0.02	0.39	1.00	11.2
	Unambiguous	94	454	9.28	0.03	0.40	0.99	12.8
	Good	6	401	8.33	0.33	0.38	0.83	12.7
	Likely	3	341	8.00	0.00	0.62	1.00	8.33

To supplement these trends, Figure 4.9 summarizes the trends in abundance, O, DBE, and C for the possible precursor molecular formulas with some evidence of artifact-formation separated by the assigned reaction pathway. All of the molecular formulas in these plots correspond to species labeled as “unambiguous”. Again, there is little notable difference between reactive pathways, especially with a much smaller set of the data. The most noteworthy observation from these plots is the staggering number of molecular formulas containing 5 O atoms in comparison to other O numbers (Fig 4.9b). It

is uncertain why this particular oxygen number is so high and may be the focus of a future investigation.

Although there is clear evidence for artifact formation during the SPE process, it is still rather ambiguous how these artifacts are forming, as there is little distinctive difference between precursors of different reaction pathways. In addition, it is difficult to know what species are responsible for the artifact with such a complex mixture.

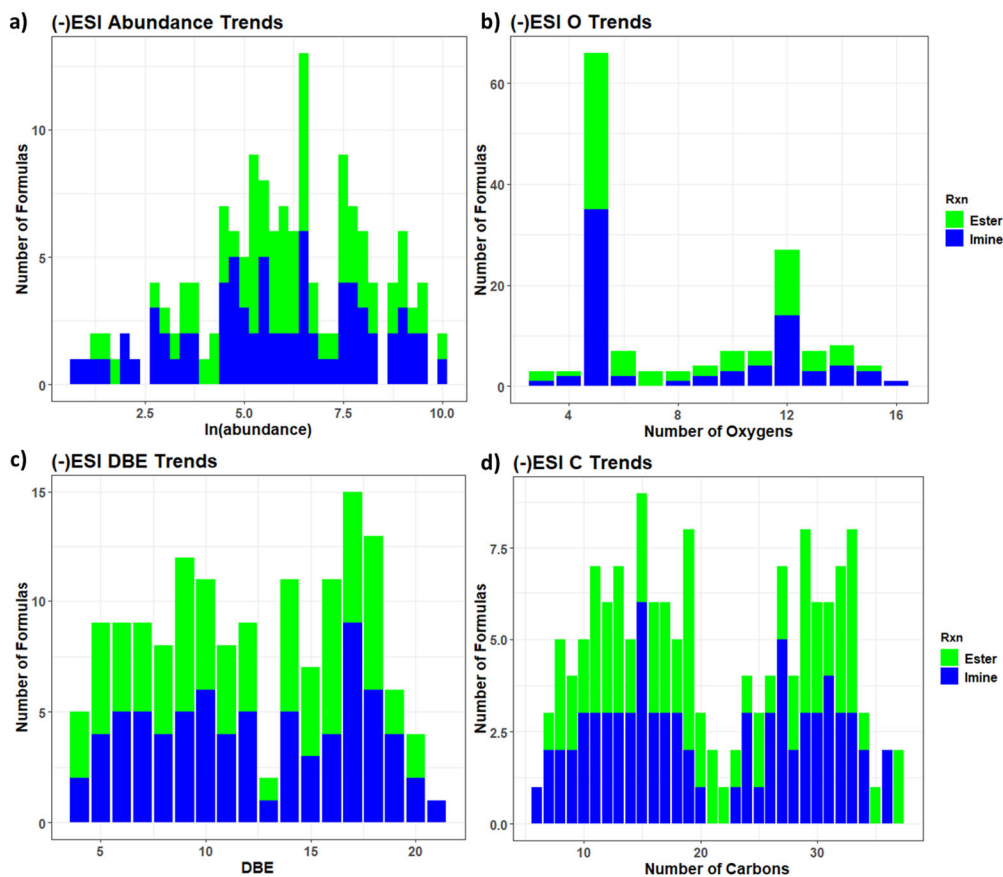


Figure 4.9. Molecular trends for potential artifact precursors in MLS with acid. All plots are colored by reactive pathway, and only includes those unique to that reaction. a) Number of formulas vs. ln(abundance); b) Number of formulas vs. number of O; c) Number of formulas vs. DBE; d) Number of formulas vs. number of C.

4.2 Positive ESI Analysis

4.2.1 Initial Molecular Formula Assignment

The data analysis for the positive ions was similar to the negative ions, however, additional intermediate steps were performed to prepare and filter the data due to the greater overall positive ion mass spectral complexity. For similar reasons, only MF less than m/z 500 were included in positive ion analysis. The goal was to determine the presence of the most likely ^{15}N artifact ions and not to attain a complete understanding of liquid smoke composition. Focusing on the lower mass range reduced data complexity inherent at higher mass ranges and potentially increased the reliability of the results, though limits the conclusions to lower mass ions. The ion distribution of L1 MLS is illustrated in Figure 4.10. HLS had similar complexity and can be assumed to have similar trends, so MLS will be the focus for the remainder of the chapter. There are large numbers of CHO and CHNO ions, as well as a significant number of CHN ions. Both CHO and CHNO compositions show a wide distribution of masses and DBE values; as a result, resolving ambiguity required additional steps for the L2 samples.

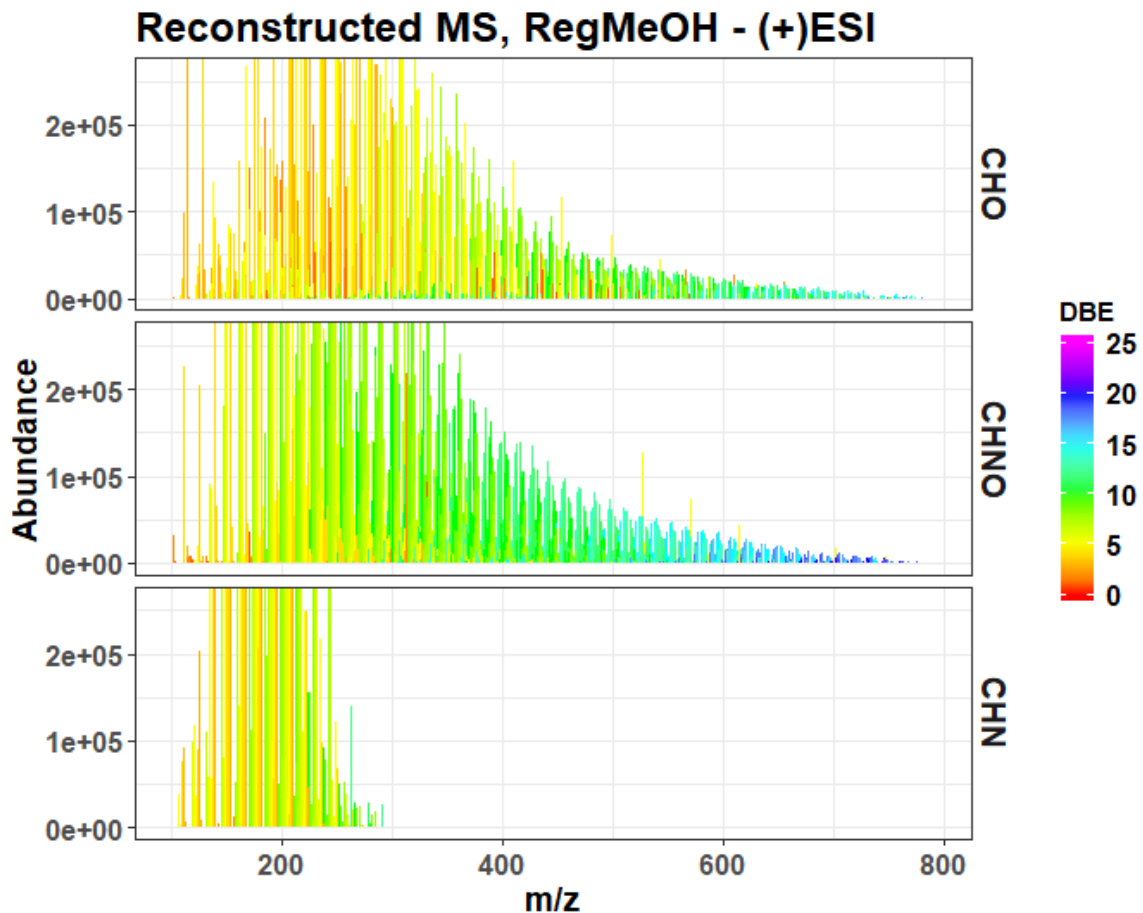


Figure 4.10. Positive mode reconstructed mass spectra for the initial assignment of the first elution (L1) MLS sample. Mass spectra are separated by group and colored by DBE.

4.2.2 Ambiguity and Quality Assurance

Similar to the negative ions, a reduction of the ambiguous MF assignments is necessary even without the ^{15}N considerations. This is due to the large number of heteroatoms and inclusion of Na^+ during molecular formula assignment, especially impacting higher m/z values. Thus, there are more chemically reasonable molecular formulas for a given measured mass within 3 ppm mass error. The chemically reasonable MF could be considered correct without applying other methods to validate them.

Ambiguity was reduced in the control samples based on two factors. First, all MF with a DBE-O (DBE minus number of oxygens) value below -10 were removed. This is based on the work of Herzsprung *et al.* (2014), which concluded MF assignments can be validated by their DBE-O values. This was based on the compositions of water-soluble dissolved organic matter, where species had DBE-O values between -10 and 10. Species with a large number of O and a very small DBE value, or highly negative DBE-O values, can be considered chemically unreasonable. In addition, due to the concentrated brown color of liquid smoke, it is inferred that these samples are very light absorbing, and thus are expected to contain large amounts of phenols, pyridines, or other aromatic moieties, as liquid smoke is a product of wood burning (Montazeri et al.2012). This implies that the lower DBE ions are less likely to be detected than higher ones for these samples.

Second, any MF with Na⁺ adducts associated with CHNO compounds were removed if the DBE > 12. This was done because Na⁺ adducts were less likely to be associated with N among the unambiguous formulas (Kruve et al. 2013). The caveat of removing those with DBE > 12 is based on the molecular characteristics of the unambiguous MF for all samples, where if a positive ion CHNO had a Na⁺ adduct, its DBE was below 13. This step is fairly arbitrary and was based only on trends in the unambiguous MF.

If a MF was still ambiguous after these steps, it was removed entirely from analysis for simplicity, since both formulas were deemed to be chemically feasible with similar numbers of heteroatoms and difficult to choose a more correct one. Less than 50 ambiguous formulas were removed as a result. The resulting molecular formulas after

these filters were then considered “unambiguous”, and were added to the original list of unambiguous assignments as the final regular NH₄OH data set.

Formulas in the ¹⁵NH₄OH samples were then filtered against the unambiguous formulas in the analogous control samples as was done previously to simplify the mass list and remove ions that are not present in the control. This also served to automatically remove most of the ambiguous ¹⁵NH₄OH MF, reducing the ambiguity from over 2500 molecular formulas to below 400 for MLS L2 without formic acid. Many of the remaining ambiguous pairs had differences in the number of heteroatoms including O, N, and ¹⁵N. Typically, the MFAssignR software takes this into account during the MF assignment and automatically removes the formula with a greater total of heteroatoms from consideration. This is based on the idea presented by Koch *et al.* (2007), where compounds with fewer heteroatoms are more likely to be chemically feasible due to rules of valence. However, since it was desirable to manually sort the ambiguous formulas due to the inclusion of ¹⁵N, this parameter was ignored until there was more confidence that crucial ¹⁵N or non-¹⁵N assignments weren't preemptively removed. However, it was deemed reasonable at this point in the filtering process to remove ambiguous formulas based on the heteroatom count. Thus, for any remaining ambiguous pairs, the one with the lowest sum of O, N, and ¹⁵N were kept. Ambiguous formulas remaining after this step (usually less than 5% the total number of formulas in the sample) were then removed completely as with the regular NH₄OH data.

Results after the filters were applied filters are shown in the elemental group separated van Krevelen plots in Figure 4.11, where Figure 4.11a summarizes the MLS L2

with acid sample and Figure 4.11b summarizes the MLS L2 without acid sample. This filter successfully removed molecular formulas outside of the control sample range to simplify the data set in both cases. Most of the higher DBE species are now absent from the plots, which was expected due to the removal of MF > m/z 500. The acidified and non-acidified samples each have similar distributions of ions, although the acidified sample appears to have a greater density of both positive CHO and CHNO ions. This is consistent with the general trend where the acidified LS samples yielded more detectable ions than the non-acidified samples, since H⁺ facilitates ionization.

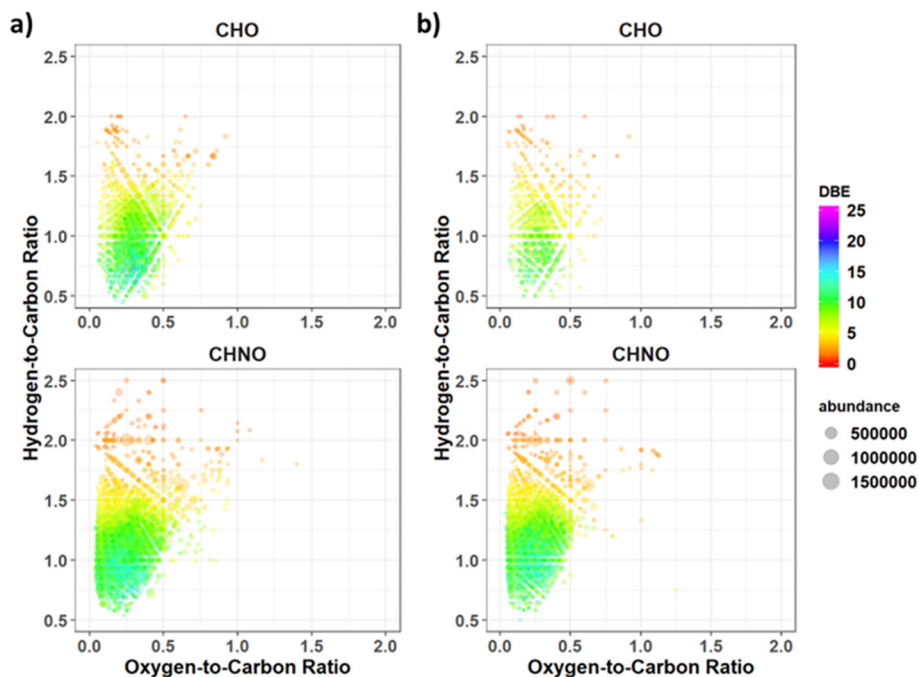


Figure 4.11. Positive mode van Krevelen plots of MLS L2 with ¹⁵NH₄OH molecular formulas after filtering. Plots are divided by group and colors indicate DBE. a) Sample without added acid; b) Sample with added acid.

4.2.3 Composition of ^{15}N Molecular Formulas

As with the negative ion mode, the absolute error plots (Figure 4.12) were used to determine the nature of the ^{15}N -containing formulas, and whether their assignments showed a similar trend as non-labeled MF or if they could be considered incorrect assignments. The MF in the plot are separated by the number of ^{15}N . In this case, the $^{15}\text{N}_0$ plot refers to any molecular formula without ^{15}N atoms, including both CHO and natural CH^{14}NO MF. There is a noticeable jump in measurement error for the $^{15}\text{N}_0$ formulas below m/z 200. With this spike in consideration, $^{15}\text{N}_1$ and $^{15}\text{N}_2$ formulas generally follow the $^{15}\text{N}_0$ error trend, indicating that both of these molecular formulas are likely the correct assignments. There are fewer $^{15}\text{N}_3$ formulas, but they also follow the general trend of $^{15}\text{N}_0$ MF. Due to these reasons, all three groups of molecular formulas containing ^{15}N were considered valid during the following assessment of the artifact composition.

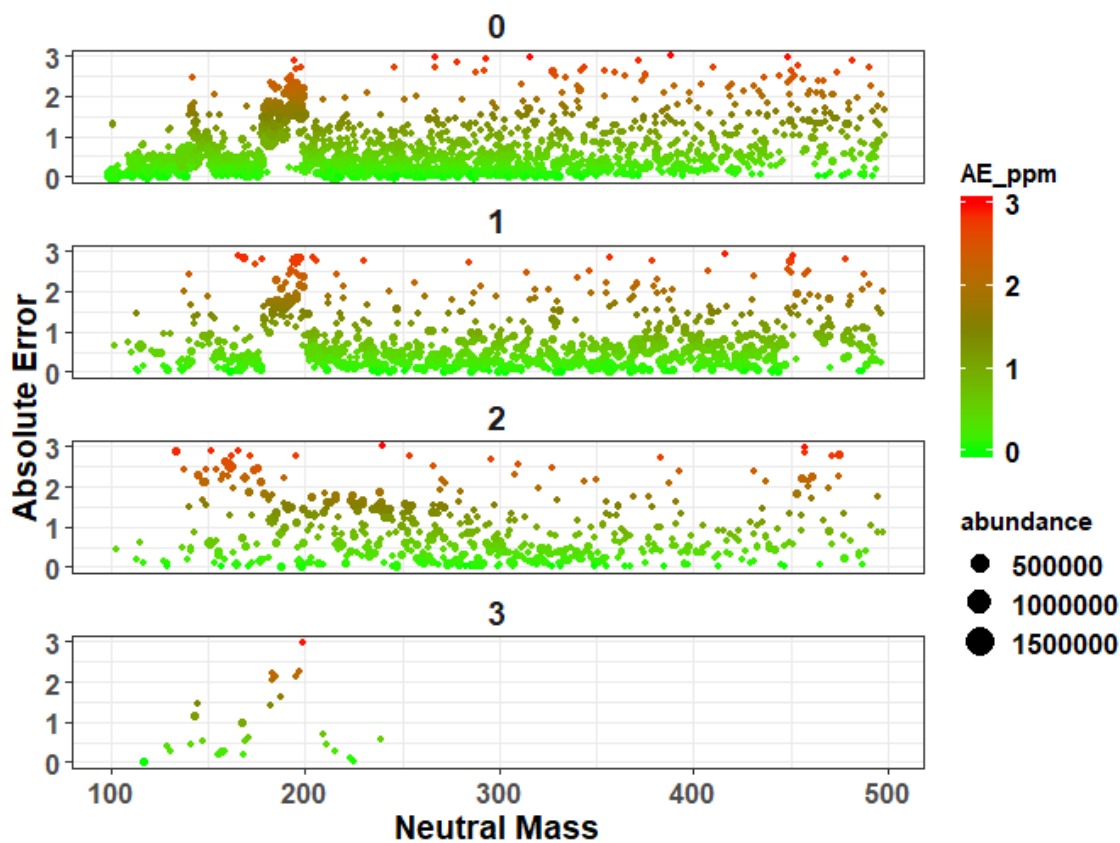


Figure 4.12. Absolute error vs. theoretical mass of MLS L2 in $^{15}\text{NH}_4\text{OH}$ with acid molecular formulas. Colors indicate error and plots are divided by formula based on the number of ^{15}N atoms assigned. Data is given in positive ESI mode.

Mass spectra for each number of ^{15}N in Figure 4.13 highlight the general ion abundance and DBE trends for each number of ^{15}N . In this case, $^{15}\text{N}_0$ refers to any molecular formula with no ^{15}N atoms. $^{15}\text{N}_1$ MF above m/z 300 have a notably higher ion abundance than those in the non-labeled MF. This could be due to artifact products with an especially high ionization efficiency, differences in the solvent composition, or CHO ions being misassigned as CH^{15}NO . $^{15}\text{N}_2$ MF were also detected at the higher m/z range, but ion abundances generally followed trends in $^{15}\text{N}_0$ where the most abundant ions were

below m/z 200. Similarly, $^{15}\text{N}_3$ MF were generally less than m/z 200, although it was difficult to describe their trends due to the low numbers and intensities of these ions. These are not shown for clarity. Overall, the compositions of MF with different numbers of ^{15}N are very different from each other.

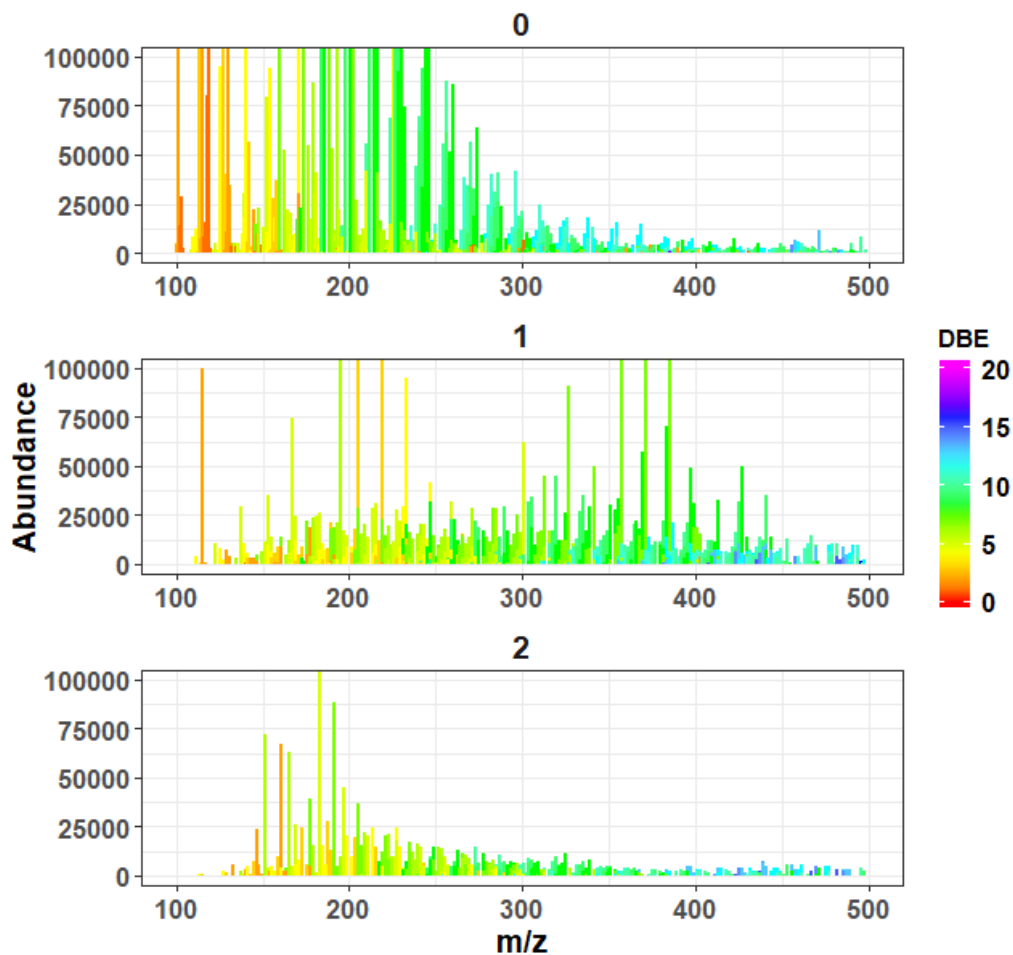


Figure 4.13. Positive mode reconstructed mass spectra of MLS L2 in $^{15}\text{NH}_4\text{OH}$ without acid molecular formulas. Colors indicate DBE and divisions are by formula based on the number of ^{15}N atoms assigned.

The number of CHNO molecular formulas for each L2 sample separated by N isotope are given in Table 4.5, which tabulates the number of formulas for each isotope of N and their percentages with respect to the entire data set. Total and average ion abundances of each are defined as the sum and average of the ion abundances of each isotopic CHNO molecular formula. Since multiple ^{15}N atoms per MF were possible for positive ESI, the ^{15}N rows includes all $^{15}\text{N}_1$, $^{15}\text{N}_2$, and $^{15}\text{N}_3$ MF. All samples had a similar or greater number of ^{15}N MF than ^{14}N with ^{15}N having up to 12.7% more molecular formulas than ^{14}N . Based on these percentages, the NH_4OH -sample interaction contributes more than one-third of the total number of MF, and at least half of all detected CHNO ions. ^{15}N MF contribute a significant amount to the total ion abundance, although this difference is not as extreme as it was for negative ESI. The average abundance for ^{15}N is also higher than ^{14}N in all samples, which is expected based on the abundances of $^{15}\text{N}_1$ ions in Figure 4.13.

Table 4.5. Number of positive ion CHNO molecular formulas of different N isotopes for all samples.

Sample Num. MF	N Isotope*	CHNO MF	%Total MF**	Total Ion Abundance ($\times 10^7$)	Avg. Ion Abundance ($\times 10^4$)
MLS L2, 0.1%FA 3571 MF	^{14}N	1244	34.8%	1.75	1.41
	^{15}N	1284	36.0%	1.46	2.37
MLS L2, 0%FA 3095 MF	^{14}N	1200	38.8%	1.44	1.20
	^{15}N	1331	43.0%	1.05	1.72
HLS L2, 0.1%FA 2957 MF	^{14}N	971	32.8%	1.09	1.12
	^{15}N	1260	42.6%	2.54	5.84
HLS L2, 0%FA 3321 MF	^{14}N	995	30.0%	0.94	0.94
	^{15}N	1419	42.7%	6.11	1.58

* ^{14}N represents CHNO MF with only ^{14}N ; ^{15}N contains MF with at least one ^{15}N .

**Calculated as a ratio of CHNO MF to Num. MF in the first column.

Various trends with respect to the number of ^{15}N formulas are given in Figure 4.14, where $^{15}\text{N}_0$ is inclusive of natural CHNO ions. The most significant difference in the distributions is in the abundance trends plot in Figure 4.14a, where there is a clear shift in the ion abundance with changing ^{15}N , as observed in Figure 4.13. Molecular formulas without ^{15}N still show a slightly lower abundance overall, but the $^{15}\text{N}_1$ MF have a unique ion abundance distribution. Trends in O, DBE, and C (Figures 4.14(b-d)) are very similar among most numbers of ^{15}N , with the exception of $^{15}\text{N}_3$ since there are fewer observed ions, which tended to center around lower DBE and C numbers.

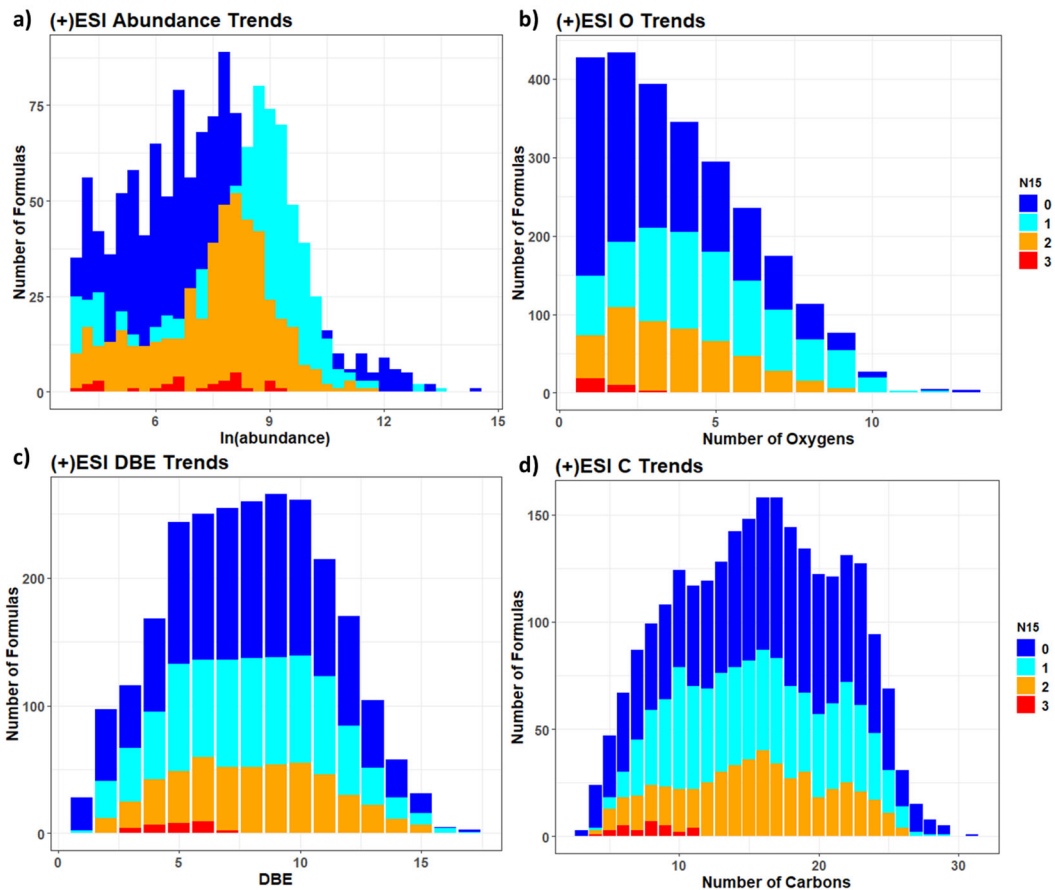


Figure 4.14. Positive mode molecular trends for MLS with acid. Colors in all plots indicate molecular formulas containing 0 or 1 ^{15}N . All plots measure number of molecular formulas vs. a) $\ln(\text{abundance})$; b) Number of O; c) DBE; d) Number of C.

4.2.4 Potential Adduct and Reaction Artifact Pathways

As discussed in Section 4.1.4, it is desirable to investigate the possible reaction pathways for the formation of ^{15}N species given the positive ion electrospray conditions. Positive ESI is an acidic process, so it was hypothesized that it may be more favorable for acid-catalyzed reactions. The reactive artifacts considered for the negative mode ions (Ester hydrolysis and imine formation) were also considered for the positive ions with two additional pathways. One additional pathway was the adduction of NH_4^+ to an

analyte during the ionization process. Since NH_4^+ is a cation, it can potentially adduct to analytes similar to H^+ or Na^+ . This would be considered an electrospray artifact rather than a reaction artifact. The second additional pathway is the formation of imidazole, a five-membered dinitrogen heterocycle. Imidazole has been found to be an important product for aqueous phase processing of isoprene secondary organic aerosol (Galloway et al. 2009; Lin et al. 2015; Hawkins et al. 2018) when in the presence of ammonium sulfate. The reaction as detailed in Scheme 2 involves aqueous NH_3 and glyoxal derivatives to produce substituted imidazoles of varying identities. Although many different substitutions are possible, the focus in this study was placed on the formation of the final product in Figure 4.15 – an imidazole with no substitution on C2.

These additional adducts and reactive artifacts were studied in a similar manner as described in Section 4.1.4 using exact mass searches for potential precursors for the ^{15}N molecular formulas. The precursors were located with a specific theoretical mass based on the theoretical mass difference between the precursor and the imidazole or adduct. Ester hydrolysis and imine formation were also considered in a similar way. The differences used for adduct and imidazole formations are summarized in Table 4.6.

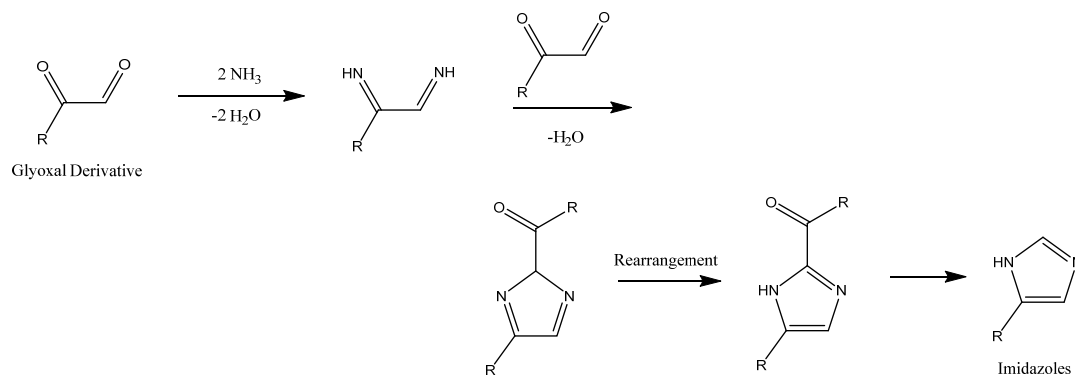


Figure 4.15. Potential mechanism for imidazole artifact formation. Based on mechanisms proposed in Galloway *et al.* (2009).

Table 4.6. Mass differences between the organic products and reactants for adduct and imidazole reactions used to locate potential precursors for ^{15}N artifacts.

Reaction Pathway	Elements Added	Elements Removed	^{14}N Mass Difference	^{15}N Mass Difference
NH_4^+ Adduct	NH_4^+	H^+	17.02655	18.02358
Imidazole Formation	N_2CH_2	O_2	10.03197	12.02604

The additional QA step performed with the negative mode ions (Section 4.1.4) was performed with the positive mode ions, where the precursor-product pair was considered valid only if the same pair was present in the control sample. This was the final QA step since it was more difficult to determine which reaction was the most likely for a given precursor. The results of these exact mass searches are summarized in Table 10, where the data for MLS and HLS without acid are given as examples. NH_4^+ adduct formation was the most likely artifact pathway based on this screening approach in both samples, followed closely by methyl ester hydrolysis. Although the theoretical masses of

all reaction precursors were very similar, NH_4^+ adduct formation occurred at a slightly lower theoretical mass on average. This is likely due to the requirement of more oxygen and subsequently a larger mass for the reaction to occur, whereas adducts can theoretically form with a wider range of compounds. The greater amount of oxygen for the reaction artifacts supports this. Imidazole precursors had the greatest amount of oxygen of all reactive artifacts, which is justifiable since the precursor requires two carbonyl groups instead of just one. As with the negative ion data, it is difficult to separate Esters and Imines by molecular parameters, with the only minor difference between the two being a marginal increase in DBE for imines. Fewer MF match to imidazole reaction precursors, which is expected since only MF with $^{15}\text{N}_2$ or $^{15}\text{N}_3$ have the potential to be imidazoles from artifact formation.

Table 4.7. Averages and counts of molecular formula parameters in MLS and HLS L2 with acid for precursors of each reaction pathway.

MLS - L2 without Acid		Averages					
Rxn	Formulas	Theoretical Mass	O	N	O/C	H/C	DBE
Adduct	457	229	3.45	0.17	0.27	0.99	7.66
Ester	438	241	3.82	0.21	0.30	1.11	7.16
Imidazole	226	243	4.39	0.18	0.37	1.10	6.94
Imine	400	239	3.96	0.20	0.31	1.06	7.37
HLS - L2 without Acid		Averages					
Rxn	Formulas	Theoretical Mass	O	N	O/C	H/C	DBE
Adduct	688	223	3.68	0.10	0.31	1.01	7.16
Ester	642	238	4.08	0.13	0.33	1.15	6.59
Imidazole	175	228	4.34	0.07	0.38	1.14	6.17
Imine	468	233	4.01	0.14	0.33	1.08	6.88

The average DBE of the adduct precursors is generally higher than that of the reaction artifact precursors, even though DBE is not a requirement for adduct formation as it is for reaction artifacts. This could indicate many of the ions that preferentially undergo adduct formation are more aromatic, as these double bonds would be inert to the reaction process. This is confirmed in Figure 4.16, where each reaction is plotted in a bar plot against the number of potential precursors for the MLS without acid sample. Colors represent the modified aromaticity index (AI_{mod}), which is a qualitative description of the aromaticity and unsaturated character of a MF based on the relative numbers of each atom proposed by Koch et al (2007). The calculation of this value was given in Equation 3.1, where in the case of the ^{15}N samples, the number of ^{15}N is included in the number of N. The number of aromatic or condensed aromatic formulas is significantly higher for the adduct formulas than those in reaction artifacts, indicating adducts are more likely to form than reactive artifacts for aromatic species. This could also be explained by the

stability of aromatic compounds. Since the reactive artifacts involve interaction with the electrophilic carbon of a carbonyl, it follows that a more reactive carbonyl will be more susceptible to attack by NH_3 . However, if this carbonyl is adjacent to an aromatic ring, the ring would be capable of stabilizing the partial positive charge on the carbon, making it less reactive and therefore less capable of undergoing reactions.

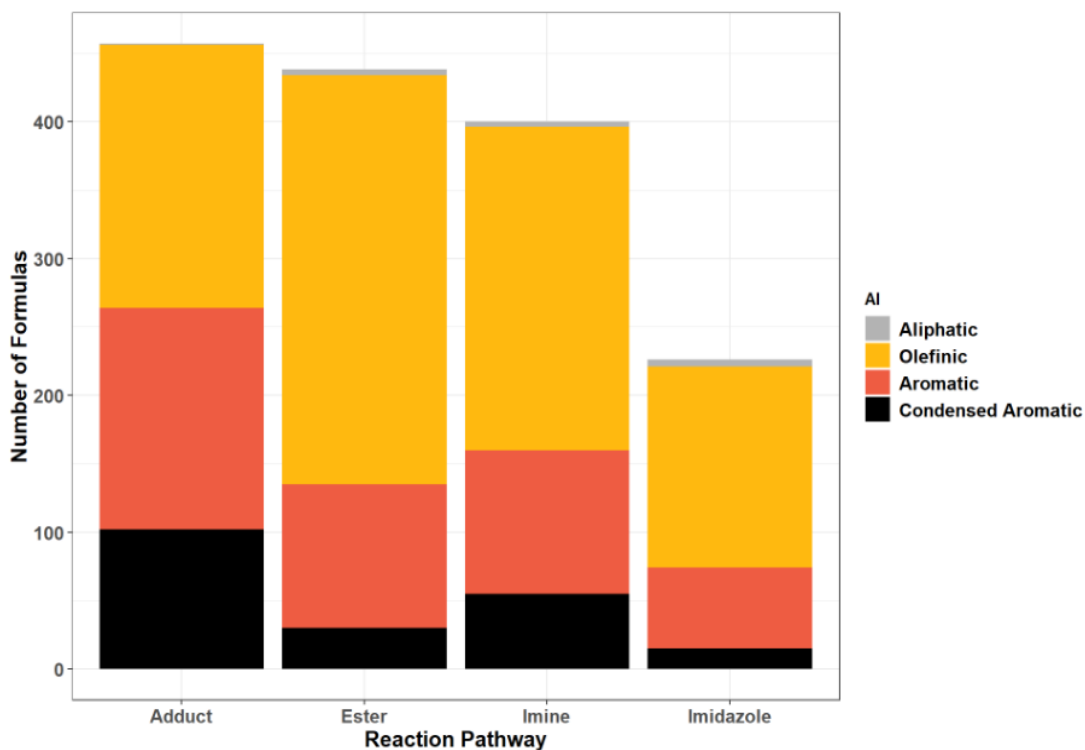


Figure 4.16. Number of formulas vs. reaction pathway in positive mode MLS L2 without acid. Colors indicate modified aromaticity index (AI_{mod}).

Elemental trends for MLS positive ions without acid are shown in Figure 4.17 and provide little distinction between the reaction types. Most, if not all, distributions among the reactions are visually similar, as might be expected based on Table 4.7. There is a

sharp decrease in the number of molecular formulas for all reactions starting with $O = 5$, then again at $O = 7$. This may be a result of the cutoff applied at m/z 500.

Although there are a large number of MF that can be traced to one or more different reactions or adducts, it is difficult to discern the true pathways for the formation of ^{15}N products. Based on methods here, there is a difference in the aromaticity between potential adduct and reaction precursors, but there is little difference between the molecular characteristics among potential reaction pathway precursors.

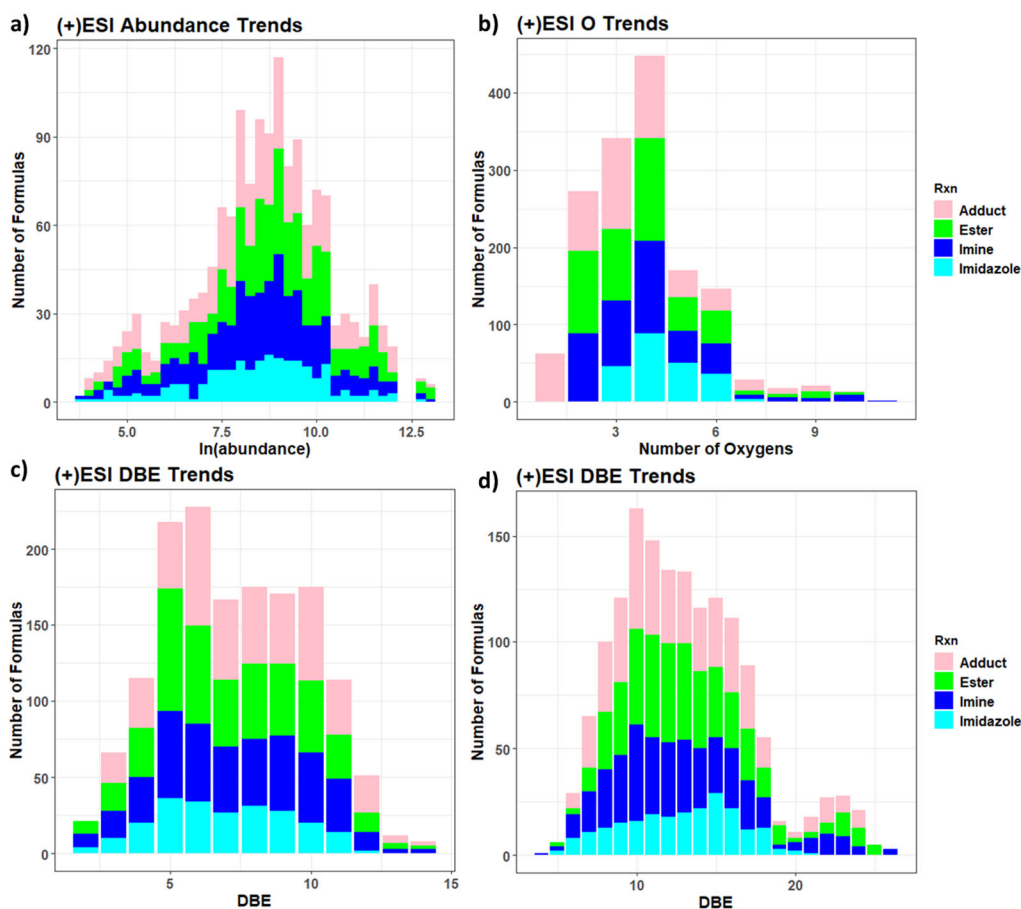


Figure 4.17. Molecular trends for potential positive mode artifact precursors in MLS with acid. All plots are colored by reactive pathway, and includes all possible precursors for each reaction. Plots measure number of molecular formulas vs. a) $\ln(\text{abundance})$; b) Number of O; c) DBE; d) Number of C.

5 Characterization of Artifacts Formed in Methanol-Solvated Liquid Smoke Using Electrospray Ionization

To connect the artifact produced analytes with a reaction pathway and determine their extent several steps are required. A preliminary assessment of the initial molecular formula composition was performed and used as a starting point for removing unlikely MF in the MeOH-d₃ sample. Then, molecular formulas with D were connected to potential reaction precursors and verified through several quality assurance steps and assumptions. Lastly, differences in the composition of the precursors for each reaction pathway were investigated to understand the molecular characteristics susceptible to artifact production.

5.1 Preliminary Molecular Formula Composition

Mass spectra for the preliminary assignment of both positive and negative ions in the regular MeOH sample are given in Figures 1a and 1b, which are separated by elemental group. CH and CHN molecular formulas are excluded due to lower numbers and they are less relevant to the reactions studied. The MeOH-d₃ mass spectra are not included due to the high ambiguity during the initial MF assignment. The abundance axes scaled to show the ion distributions at higher m/z values, and the abundance values are 10 times higher in the positive mode than negative mode due to the greater ion intensities. This is related to the amount of reduced N in biomass burning samples (Laskin et al. 2009). Ions are well distributed past m/z 700 in both ionization modes with increasing DBE for increasing mass. The bimodal distribution in negative CHO ions is typical for

biomass burning samples and is a result of oligomerization due to monoterpene influences (Zhao et al. 2013). As expected, there is a greater number of positive CHNO ions than negative CHNO ions due to the N basicity. The tall ions around 300 are S-containing contaminants present in MeOH and were removed in future processing steps. Also, there is a noticeably high density of positive CHNO ions up to m/z 600, which are likely polynitrogen species (Galloway et al 2009; Hawkins et al. 2018). Generally, the positive ion mass spectra are more complex than the negative ion spectra, due primarily to the larger number of ions and the possibility of multiple adduct ions (H^+ and Na^+).

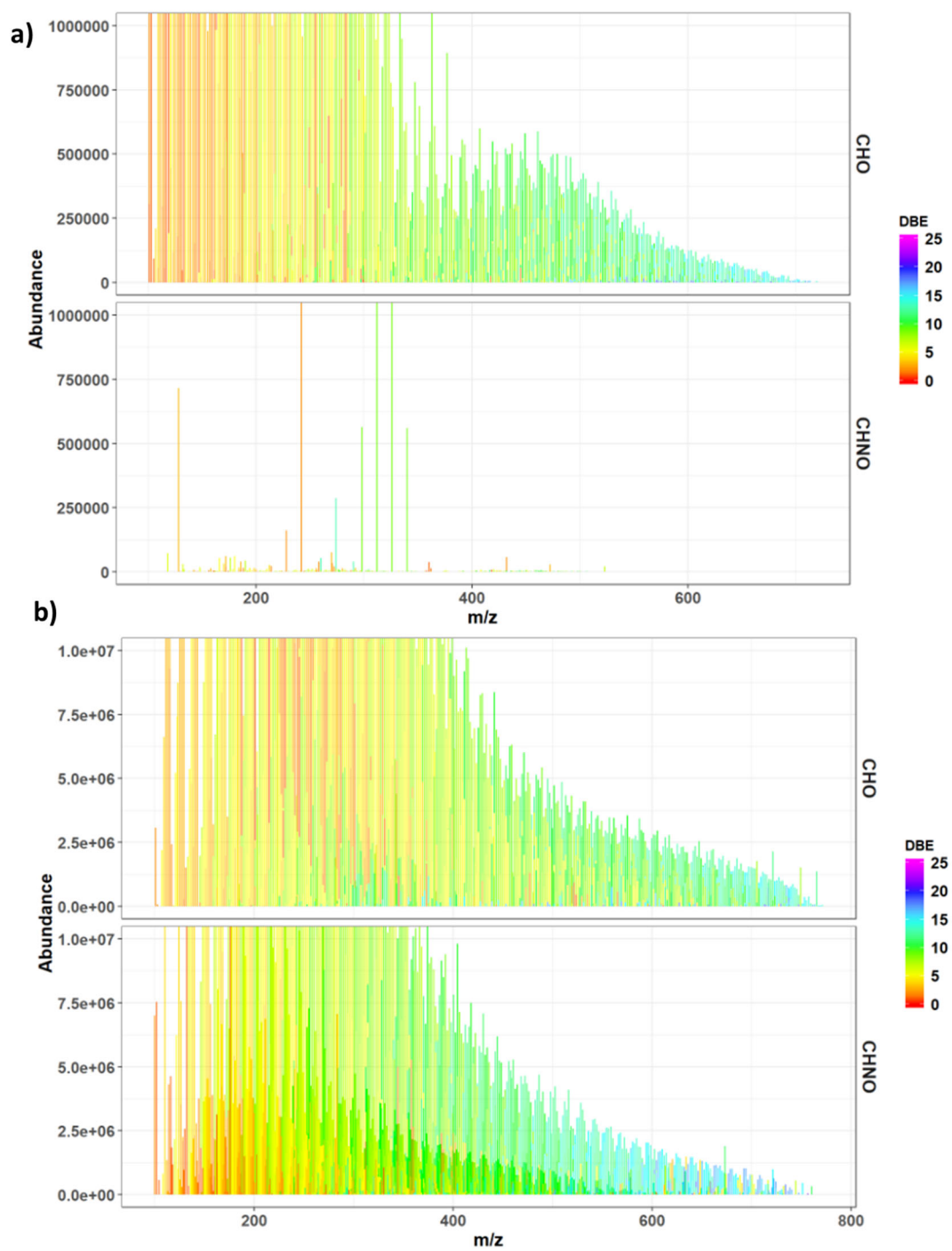


Figure 5.1. Reconstructed mass spectra for MLS in a) negative ion mode, and b) positive ion mode. Ions are divided by elemental group and colors represent DBE.

van Krevelen plots for the negative (Figure 5.2a) and positive (Figure 5.2b) ion molecular formulas highlight the complexity of the MeOH-d₃ MF assignment. The figures are divided by elemental group and solvent. The regular MeOH samples have van Krevelen distributions expected for biomass burning complex mixtures (Brege et al. 2019), with assignments generally following a trend line of increasing H/C for increasing O/C and the highest DBE values at H/C < 1 and O/C < 0.5. This indicates the samples have many aromatic constituents, which is expected since the liquid smoke samples are expected to have many phenolic compounds (Montazeri et al. 2012). The positive CHNO ions are more complex with many assignments in the aliphatic region of the plot. MeOH-d₃ plots are additionally more complex than MeOH with unusually high densities of ions in many regions, such as the cluster centered around 1.3 H/C and 0.5 O/C for positive CHNO ions. This is due to the ambiguous assignments, which were removed in future filtering steps.

A comparison of MF assignments before reducing the ambiguity are given in Table 5.1. The data is divided by ion polarity and solvent. There are generally more MF assignments of positive ions even with ambiguity considered, which is consistent with Figures 5.1 and 5.2. More MF were also assigned in MeOH-d₃ for both polarities due to ambiguity. However, fewer positive CHNO ions were assigned in MeOH-d₃ than regular MeOH, potentially due to CHNO ions with D₃ being misassigned as CHO. There is a low number of CHN and CH assignments in both polarities (< 2% their respective totals), and although are likely correct assignments, they were omitted from the analysis.

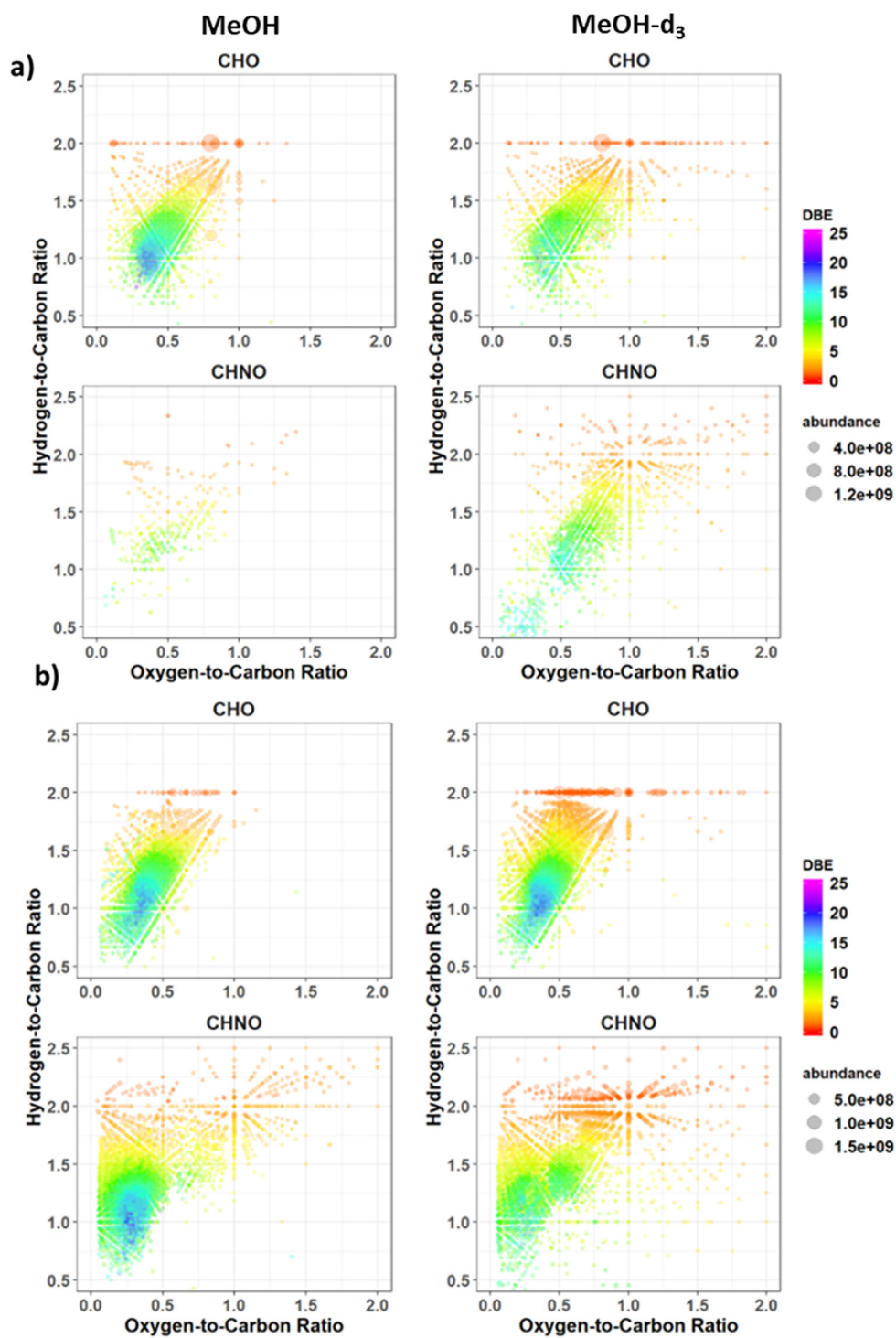


Figure 5.2. Initial van Krevelen plots for a) negative mode ions, and b) positive mode ions. Plots are divided by elemental group for each sample and colors represent DBE. The leftmost plots in both a) and b) represent ions in regular MeOH, and the rightmost plots represent ions in MeOH-d₃.

Table 5.1. Ambiguous and unambiguous molecular formula for each elemental group in each solvent with respect to the ionization polarity.

Group	Negative Ions		Positive Ions	
	MeOH	MeOH-d ₃	MeOH	MeOH-d ₃
CHO	1743	2083	2164	2781
CHNO	277	1509	3623	3050
CHN	1	10	74	49
CH	1	7	13	10
Total	2022	3609	5874	5890

5.2 Preliminary Filtering of Formulas

As with the NH₄⁺ artifact, molecular formulas were filtered to simplify the data set and limit the analysis to only the most likely molecular formulas. Although several filters were applied to the data, this section discusses two preliminary filters to refine the data set and explore molecular characteristics.

The first filter applied to the MeOH-d₃ sample was to remove MF with unexpected numbers of D. The reactions of interest would include the per-deuterated methyl group and thus the most likely artifact molecular formulas should only contain D in multiples of 3. Therefore, only MF containing 0, 3, or 6 D were considered as potentially valid. For the second filter, MeOH-d₃ MF were compared to the control formulas as described in Section 2.3.2. This filter removes all of the remaining ambiguity from both polarities of MeOH-d₃ ions. These filters greatly reduced the complexity of the MeOH-d₃ data (Figure 5.3) making them more comparable to the regular MeOH data, as expected.

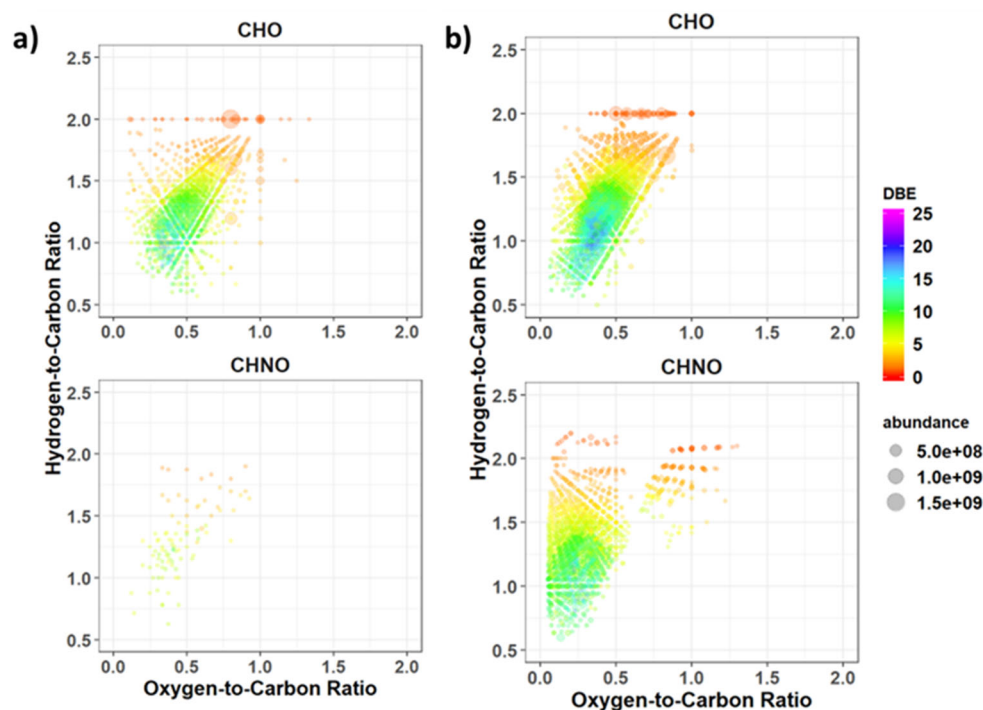


Figure 5.3. van Krevelen plots of MeOH-d₃ after filtering against regular MeOH in a) negative mode; b) positive mode. Plots are divided by group and colors indicate DBE.

The improved reconstructed mass spectra post-filtering are given in Figure 5.4a and 5.4b representing negative and positive ions, respectively. The mass spectra are divided by the number of D in the assigned MF. Both polarities display lower deuterated ion densities. These MF are also generally lower in intensity, which is to be expected based on the nature of the reactions of interest, as they typically require a strong acid to proceed (Reusch 1999). While there are few noticeable trends for the negative ions, there is a marked overall reduction in the DBE values for the positive deuterated ions, providing evidence for a reaction artifact. A significant difference between the positive and negative ions is due to the presence of D₆ ions, which were absent in the negative

ions. This implies the electrospray process may have some impact in shifting the direction of the equilibrium for the artifact reactions.

The numbers of MF, total abundance, and average abundance for each ionization mode are summarized in Table 5.2. Ion type is separated by the number of D, and the %total MF was calculated relative to the total number of MF in that ionization mode. The TIC (Total ion current) represents the sum of the total abundance. About 15-20% of the MF in both ionization modes are deuterated with a slightly larger percentage for negative ions. However, positive D₃ and D₆ ions compose a larger amount of the positive TIC than negative D₃ ions do for the negative mode. In addition, the average abundance of deuterated ions is more significant for positive mode. This indicates that even though there are more D formulas in negative mode, they may have more impact in positive mode. However, these are raw numbers of D-containing formulas; and these numbers were reduced after connecting MF to potential reaction precursors.

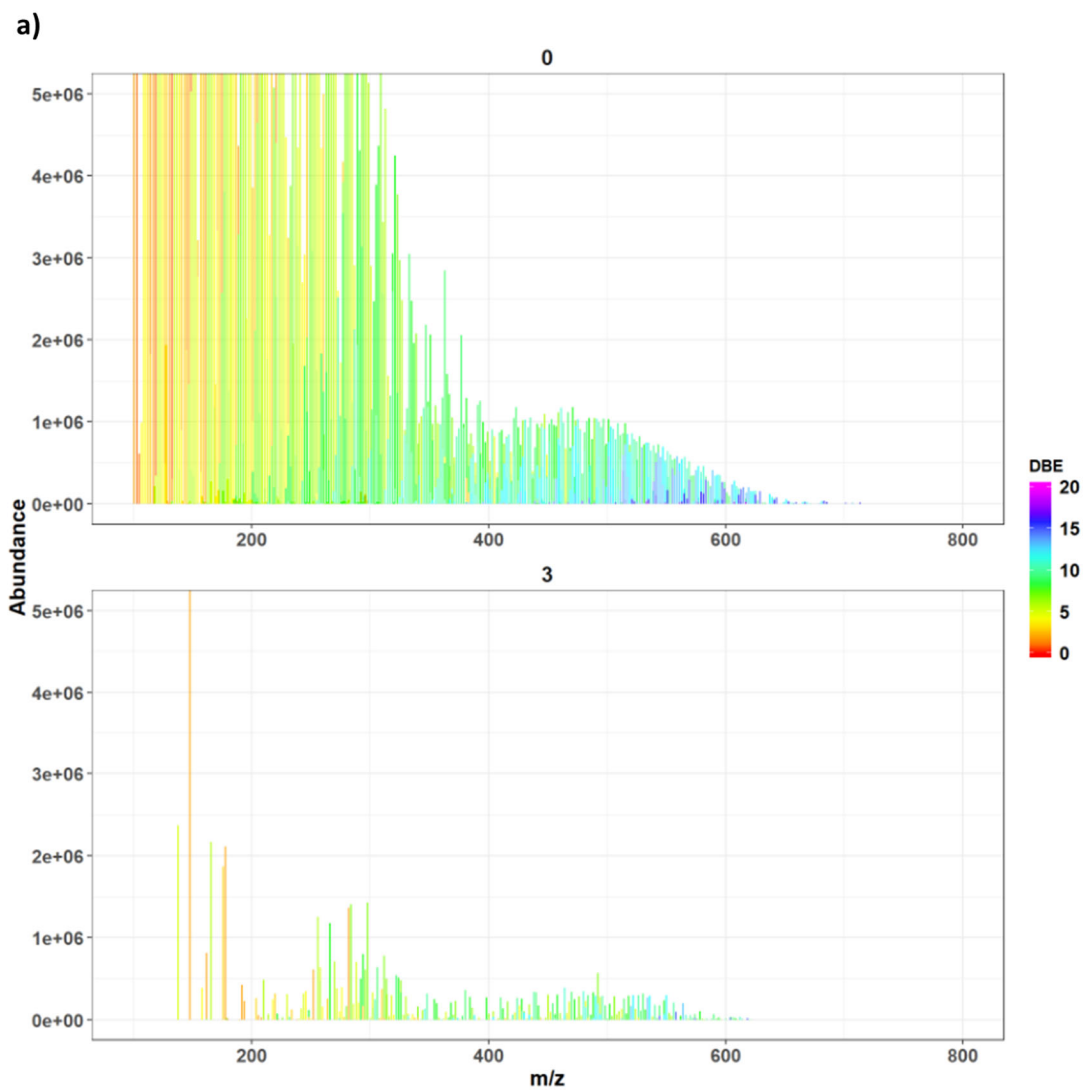


Figure 5.4. Reconstructed mass spectra for filtered MeOH-d₃ data obtained in a) negative mode; b) positive mode. Molecular formulas are divided by number of D, and colors represent DBE.

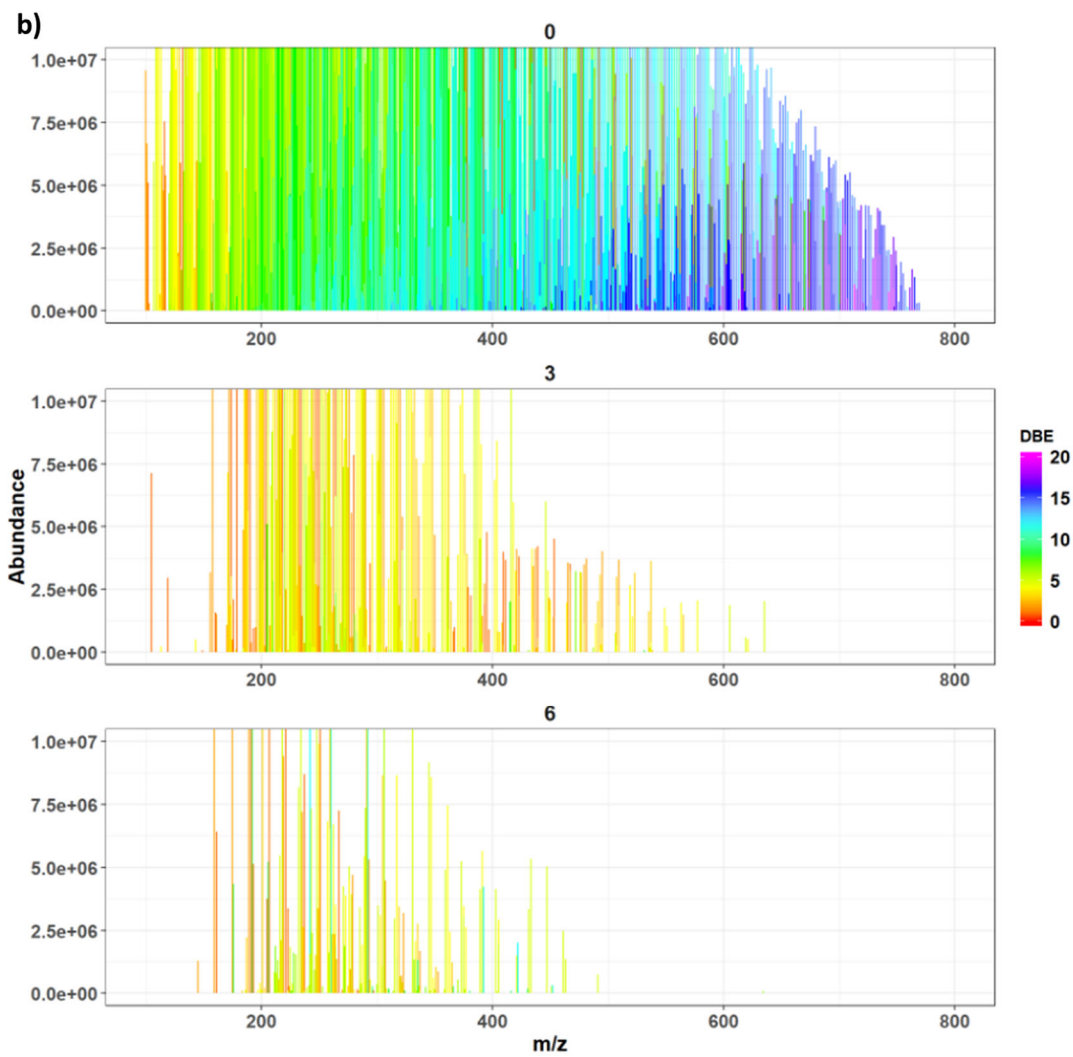


Figure 5.4 (cont.) Reconstructed mass spectra for filtered MeOH-d₃ data obtained in a) negative mode; b) positive mode. Molecular formulas are divided by number of D, and colors represent DBE.

Table 5.2. Number of molecular formulas with each number of D for negative and positive ions.

Ionization Mode Total MF	Number of D	MF	%Total MF*	Total Ion Abundance (x10⁸)	Avg. Ion Abundance (x10⁵)
Negative Ions 1567 MF	0	1266	80.8%	88.23	69.69
	3	301	19.2%	0.78	2.60
Positive Ions 3542 MF	0	2960	83.6%	630.88	213.14
	3	386	10.9%	59.07	153.02
	6	196	5.5%	11.10	56.63

*Calculated based on Total MF in the first column.

Various molecular characteristics for the positive ions are given in Figure 5.5, where the number of MF with respect to abundance, number of O, DBE, and number of C are plotted for D₀, D₃, and D₆ ions. Negative ions are not included as the D₃ trends closely matched the trends of D₀ and did not offer new information on the molecular characteristics in the sample. The positive ion D₃ and D₆ MF follow the general abundance trend of the positive D₀ MF, where there is a large variability in abundance (Figure 5.5a). The D₆ MF are shifted to a lower abundance than the D₃ MF, which is expected based on the reaction pathway. Oxygen numbers (Figure 5.5b) tended to be lower in the deuterated MF than in non-deuterated MF, although they had roughly equivalent modes around 4-5 O. Also there is a small, secondary mode for D₃ ions. This could potentially be due to the characteristics of the sample, although this feature was more prominent for negative ions. Carbon numbers (Figure 5.5c) were similar to the oxygen numbers, although the primary mode was around 12 C and the secondary mode was not present. Also, D₃ formulas were not measured at higher carbon numbers unlike they were with higher oxygen. However, there were a few outlier D₃ and D₆ MF around

24 C, which cannot be explained by the bimodal distribution. The cause of these outliers is unclear.

DBE values (Figure 5.5d) show a more distinct trend for deuterated ions. All D₃ ions had DBE values < 9 and a majority of D₆ formulas showed this trend as well with the exception of outliers above 10. Although the shape of the D distributions is roughly similar to the D₀ distribution, it is possible that D may not interact with the sample species at higher DBE values. This could indicate that these reactions may not occur in the presence of highly unsaturated species, such as aromatic systems. This will be explored further as these MF are evaluated for their potential reaction pathways.

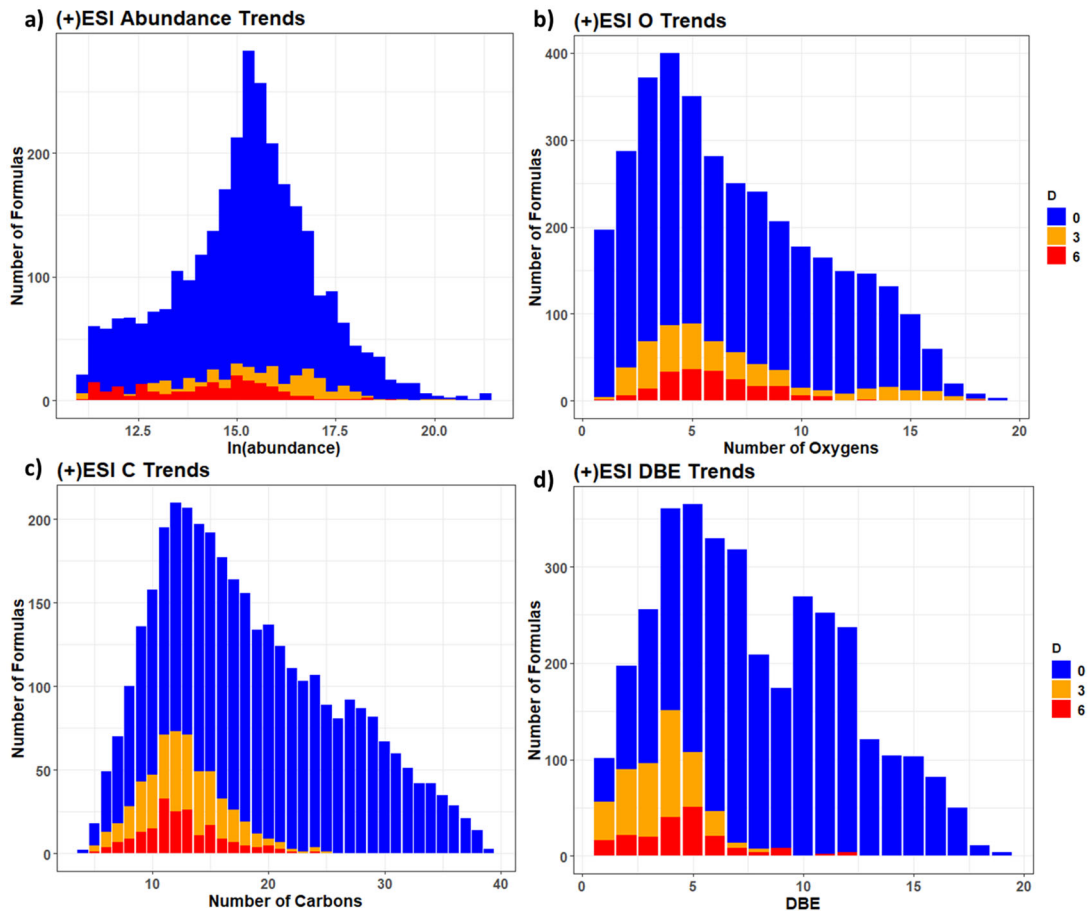


Figure 5.5. Molecular trends for positive mode MeOH-d₃. Colors in all plots indicate the number of D in each molecular formula. All plots measure number of molecular formulas vs. a) ln(abundance); b) Number of O; c) Number of C; d) DBE.

5.3 Potential Pathways of Artifact Formation

With the evidence of the MeOH-d₃ interactions with the MLS analytes, the next step was to determine if these artifacts were a result of certain reaction pathways, and if so, what molecular characteristics were most likely to promote this artifact formation. The focus of this investigation was restricted to reactions with carbonyls due to their prevalence in biomass burning processes and by extension MLS.

As with the NH_4^+ artifact formation, these reactions are expected to occur predominantly in acidic conditions and have the potential of occurring in solution. However, the ESI process can alter the pH of the solution by inducing an excess charge in solution, which concentrates the droplets (Zhou et al. 2002). Therefore, it is possible that positive ESI may accelerate this reaction. As such, it was necessary to study these reactions in both polarities to understand the dominant factor in controlling the artifact formation.

Three related reactions were studied in positive ESI based on the studies of Bateman *et al.* (2008). All methanol reactions involve a carbonyl group with either acyl addition or substitution, and thus all steps are in equilibrium (Figure 5.6). Reaction A involves a molecule of methanol reacting with a carboxylic acid via nucleophilic acyl substitution to eliminate water and form a methyl ester. The results from this reaction are subsequently labeled “Ester”. Reaction B is the nucleophilic addition of a molecule of methanol across the carbonyl of a ketone or aldehyde. This results in a hemiacetal product, where an ether and alcohol are bound to the same carbon. The results from this reaction are subsequently labeled “Hemi”. Reaction C is a continuation of reaction B, where the hemiacetal product reacts with another molecule of methanol to eliminate water and form an acetal, with two ethers bound to the same carbon. Since it requires a hemiacetal to be present, reaction C should be less likely than A and B. Reaction C will be labeled “Acetal” for the remainder of this work. Only reactions A and B were assessed for the negative ions since no D_6 MF passed the filtering steps. In addition, only negative CHO ions were investigated since there were few negative CHNO ions.

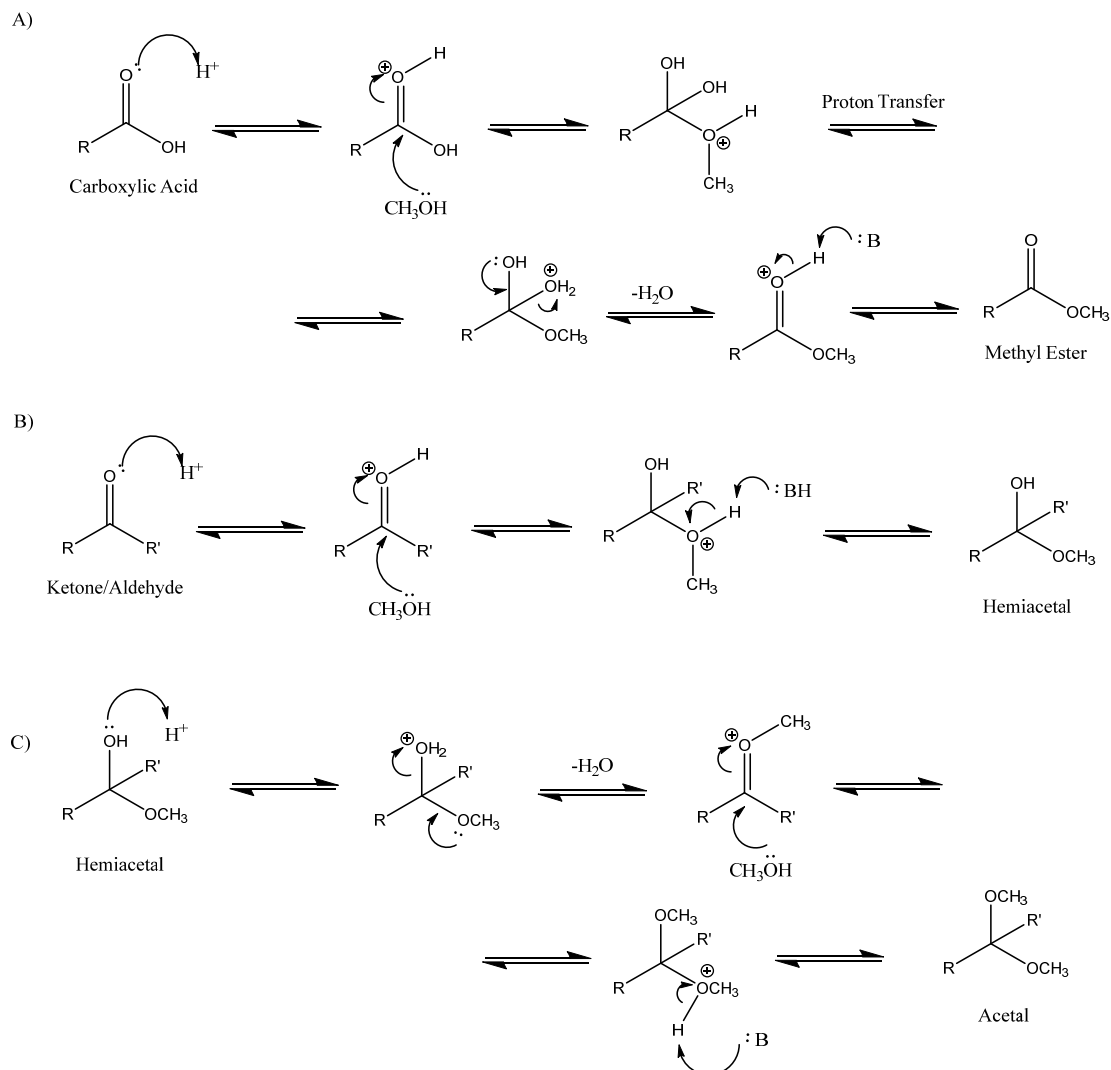


Figure 5.6. Proposed mechanisms of potential reactions between MeOH and carbonyls leading to deuterated formulas. A) Fischer Esterification from a carboxylic acid; B) Hemiacetal formation from a ketone/aldehyde; C) Acetal formation from a hemiacetal formed in B).

To determine if these reactions were potentially occurring, all filtered D_3 and D_6 molecular formulas in MeOH- d_3 were screened against D_0 formulas to determine if the precursors existed for these reactions based on theoretical mass differences, as described

in Section 4.1.4 for ^{15}N molecular formulas. Regular MeOH MF were treated similarly, but the D_0 mass difference was used instead of the deuterated mass difference. The mass differences for D_0 and D_3 (or D_6) are summarized in Table 5.3. Reactions A and B were evaluated for D_3 MF, and Reaction C was evaluated for D_6 MF, since acetals would theoretically add two CD_3OH groups to the precursor.

Once the matches for each reaction were found in MeOH- d_3 , these matches were again filtered by applying three additional criteria. First, any artifact-precursor pair in MeOH- d_3 was required to exist in regular MeOH, as was done with the $^{15}\text{NH}_4\text{OH}$ samples prior. Second, the relative abundances of each reaction ion were evaluated against the potential precursor, and only those with a product ion abundance less than that of the precursor were kept. Since these reactions require sufficiently acidic conditions to promote the equilibrium to shift, it was assumed that the precursor should be more likely to exist than the produce. Most precursor/product pairs displayed this trend. Third, acetal precursors were retained only if the same Hemi precursor was present. Since acetal formation is an equilibrium process, this reaction was not expected to occur without the hemiacetal intermediate also being detected. (Reusch 1999; Bateman et al. 2008)

Of the 386 positive D_3 MF, 272 had a potential Hemi precursor, while 229 had a potential Ester precursor. In addition, 66 of the 196 positive D_6 MF had a potential Acetal precursor. The numbers were similar for the negative ions, although no D_6 MF existed. There was a large amount of overlap between Ester and Hemi MF, indicating these molecular formulas could have resulted from both reaction pathways.

Table 5.3. Mass differences used to locate potential precursors for each reaction in regular and MeOH-d₃.

Reaction Pathway	Elements Added	Elements Removed	MeOH Mass Difference	MeOH-d ₃ Mass Difference
Ester	CH ₃ OH	H ₂ O	14.01565	17.03448
Hemi	CH ₃ OH	N/A	32.02621	35.04504
Acetal	2 CH ₃ OH	H ₂ O	46.04186	52.07952

Due to the amount of overlap, it was desirable to determine the reaction pathway most likely to lead to the production of an artifact ion for those with both possible precursors. This was performed similarly to the ¹⁵NH₄OH artifact evaluation (Section 4.1.4). Briefly, MeOH-d₃ ions were assigned to reactive precursors based on how closely each reaction's precursor/product ratios matched to the same precursor/product pair in the regular solvent. Labels were assigned based on the percent difference between each reaction to denote their likelihood. All acetal formulas were automatically unambiguous, since D₆ MF were only evaluated for acetal precursors. It is possible there are D₆ ions resulting from two successive methanol reactions (either Reaction A or Reaction B in Figure 5.6) in the same molecule, however, these reaction combinations were not studied.

Compared to unambiguous MF, there were very few reaction artifacts with *Good* or *Likely* classifications. As a result, only the unambiguous assignments for each reaction are provided in Table 5.4, which tabulates the averages of several molecular characteristics in both negative and positive mode. This was additionally done to investigate the most different reaction precursors. The column "Rxn % RA" refers to the

average percentage of the deuterated, post-reaction ion abundance compared to the potential precursor abundance. In the negative mode, there were over three times more unambiguous Ester reactions than Hemi reactions, although both had relatively low numbers. On average, negative mode Hemi reactions had a greater precursor mass, number of O, and DBE. This indicates the precursors for negative mode tended to be more oxygenated, unsaturated, and had higher mass. For the positive mode, the trends in the number of MF, mass, and oxygen content are reversed. Ester reactions had fewer MF, were typically larger, and had more oxygen. The major difference was the DBE value, where the average DBE for a species capable of undergoing an esterification reaction was less than 3. The DBE values for Hemi and Acetal reactions were also relatively low compared to negative mode ions with averages just greater than 4. Acetal precursors had slightly lower average mass than Hemi precursors, but the O/C was slightly higher for Acetal precursors. This indicates Acetal reactions are more likely to occur with smaller, oxygen-rich species.

The relative abundances of the reaction artifacts compared to the precursors for each reaction in positive mode were noticeably different. As expected, the Acetal reactions have lower average percent relative abundance than Hemi reactions, as Acetal reactions should only occur to a small extent with formed hemiacetals. In comparing the ester formation to the hemiacetal formation however, it appears the Ester reactions result in more abundant products than Hemi reactions, even though there are fewer unambiguous Ester precursors. This could indicate that there are fewer species in MLS capable of undergoing esterification, but when they do, the equilibrium is driven more

towards the ester product. The reason for this requires further investigation. Although the overall percentages differ between the reactions, they are likely not extremely significant based on the TIC in Table 5.2, as positive D-containing ions only represent about 10% of the total abundance. In addition, the ions connected to a reactive pathway would make an even smaller percentage of this total abundance. For a large majority of the post-reaction MF in MeOH-d₃, unambiguous or not, the same product MF in the control sample had a larger % RA (not shown). This implies that even if it is assumed that some of the post-reaction MF did result from a reaction, there is still an amount of the ion abundance that did not, indicating it was natural in the sample. Thus, although the reaction may occur to some extent, it is not severe, as the reactive product MF is already present in the sample.

Table 5.4. Molecular formula parameter counts and averages for potential reaction precursors defined as *Unambiguous*.

Negative Mode		Averages - Unambiguous Formulas						
Rxn	MF	Precursor Mass	O	N	O/C	H/C	DBE	Rxn % RA
Ester	51	363	8.55	0.00	0.51	1.25	7.51	24.30
Hemi	15	438	10.70	0.00	0.55	1.31	8.20	25.44
Positive Mode		Averages - Unambiguous Formulas						
Rxn	MF	Precursor mass	O	N	O/C	H/C	DBE	Rxn % RA
Ester	45	330	9.04	0.51	0.73	1.73	2.84	33.90
Hemi	88	262	5.50	0.75	0.48	1.49	4.35	20.60
Acetal	68	252	5.40	0.37	0.50	1.40	4.49	6.86

Based on the averages in Table 5.4, it seemed molecular formulas capable of undergoing reactive artifacts had relatively low DBE values. To investigate this further, all D-containing molecular formulas in MeOH-d₃ were plotted against the non-deuterated

formulas in the van Krevelen space for the negative ions (Figure 5.6a) and the positive ions (Figure 5.6b). The plots are divided by elemental group. The gray points indicate the non-deuterated MF in the MeOH-d₃ sample, while the colored points indicate D-containing MF in the same sample, where the colors indicate the DBE value. Notably, negative D-containing MF are restricted to CHO ions due to the low number of CHNO MF and positive D-containing MF are present in for both CHO and CHNO. Negative D₃ ions appear to have no discernable pattern with the MF present in the center of the D₀ MF and with varied DBE values. However, most MF had DBE values below 15. Positive D-containing ions showed a more distinct pattern with most ions above an H/C ratio of 1.5 and low DBE values for both CHO and CHNO ions. This is consistent with the low DBE values described in Table 4 and implies many of these MF are more aliphatic than aromatic.

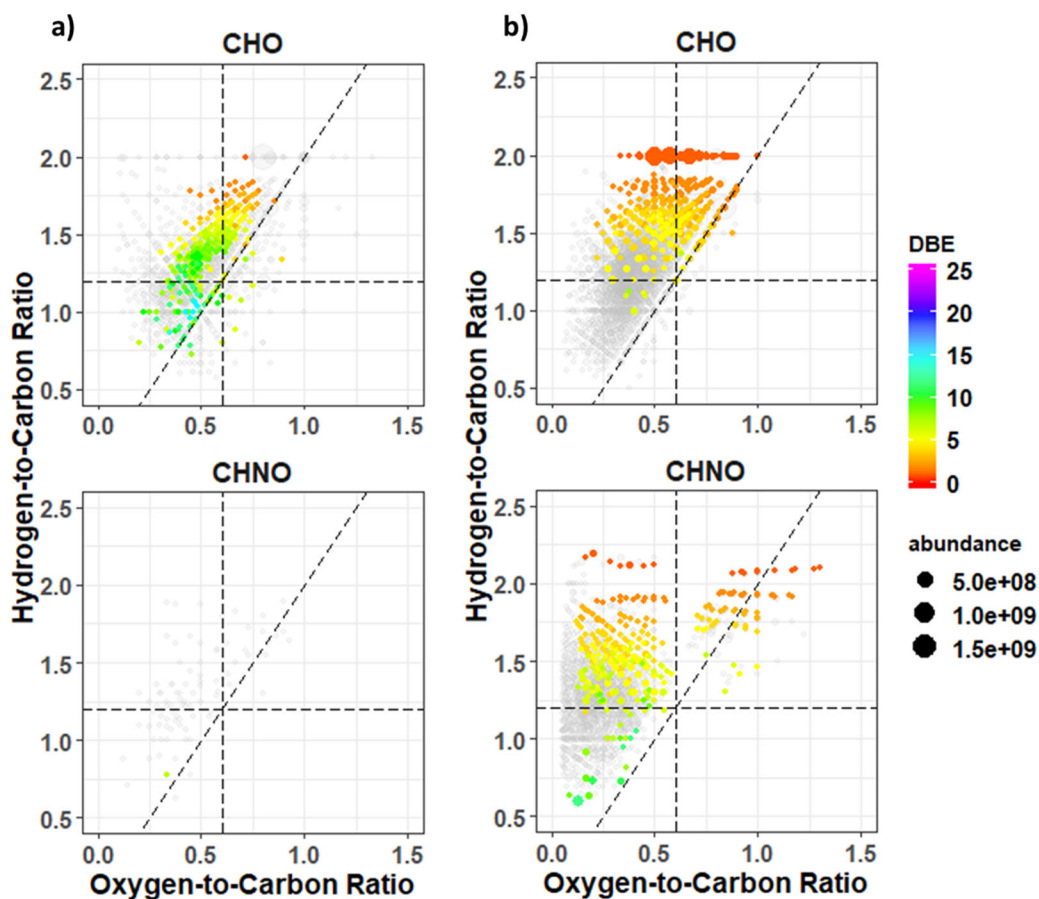


Figure 5.7. van Krevelen plots of deuterated ions versus non-deuterated ions in a) negative ion MeOH-d₃ and b) positive ion MeOH-d₃. Plots are divided by elemental group. Gray points represent non-deuterated formulas, while points in color represent deuterated (D₃ or D₆) ions, where colors indicate DBE.

The histograms of the MF aromaticity for all positive and negative MeOH-d₃ MF are summarized in Figure 5.7. All vertical axes are the number of molecular formulas and colors indicate the calculated modified aromaticity index (AI_{mod}), as described in Section 3.2.3. The histograms indicate either the number of D in the MF (Figure 5.7a and 5.7c), or the reaction pathway (Figure 5.7b and 5.7d). The reaction pathway includes the AI_{mod} of the reaction precursors and not the reaction products. Although there was little trend in

the DBE values for the negative ions, it is clear the deuterated MF are generally non-aromatic, even though there is evidence of aromatic and even condensed aromatic ions for non-deuterated MF. The same is true for positive ions; in fact, most D-containing positive ions were calculated to be aliphatic. Few of the reaction precursors in each mode were classified as aromatic and only Hemi and Acetal reactions included aromatic MF for positive ions. This indicates that these reaction artifacts may be more likely to occur with aliphatic molecules. In other words, aromatic compounds are less reactive under these conditions. This has significant implications since MLS and biomass burning samples are rich in phenolic products and other aromatic compounds, and thus they are less susceptible to methanol reaction artifact formation.

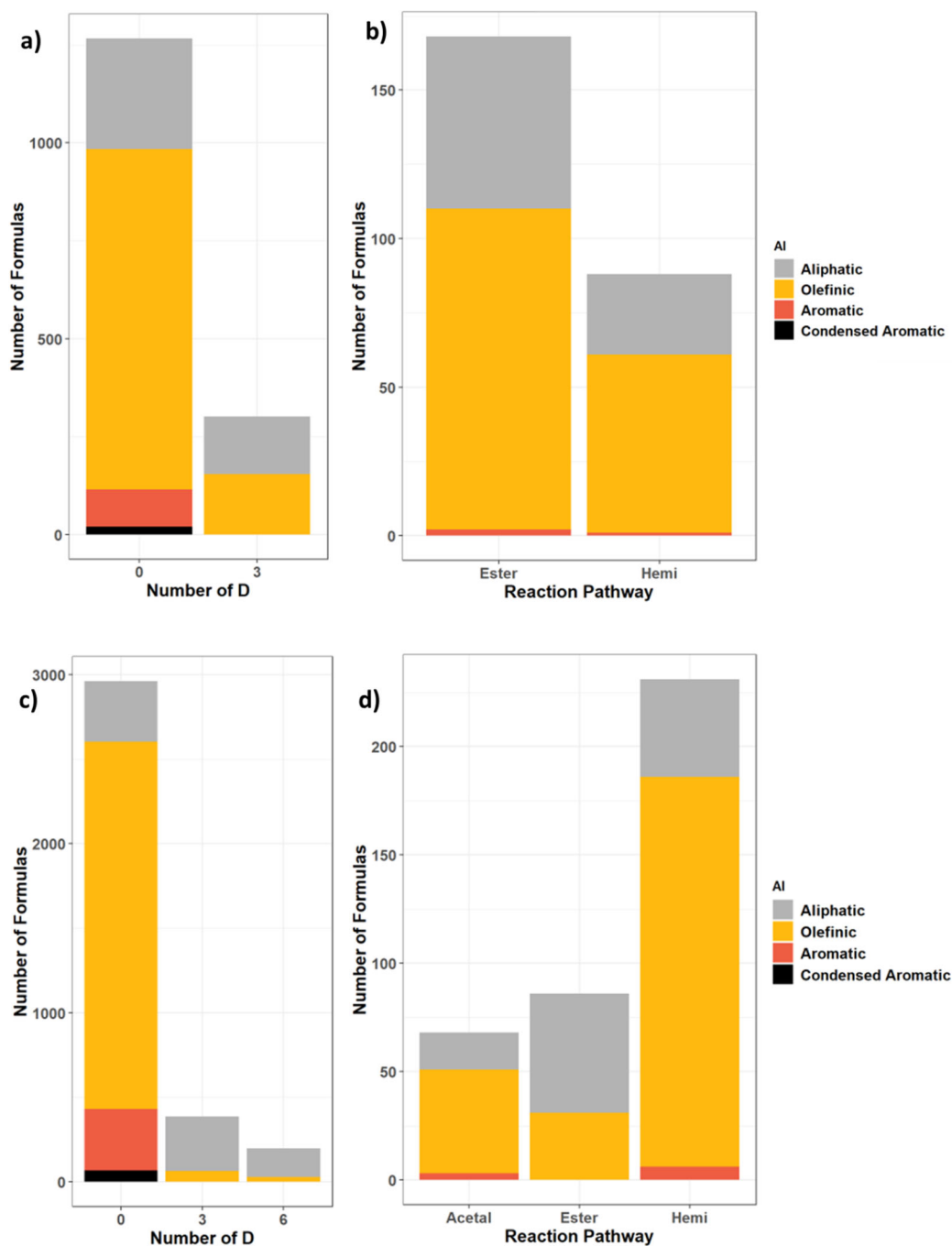


Figure 5.8. Aromaticity plots for each ionization mode, with bars colored by calculated aromaticity index (Koch et al. 2007). Figures 5.8a and b) display negative ions, while c) and d) display positive ions. Figures 5.8a and c) represent all assigned MF in MeOH-d₃, grouped by number of D, while Figures 5.8b and d) represent only the unambiguous precursors of MF assigned to a reaction pathway, separated by reaction.

6 Conclusions

Evidence that sample preparation and the electrospray ionization process can change the detected composition of biomass burning samples was observed using an MS informatics approach and new tools. It is especially necessary to use caution when adding NH_4OH , because NH_3 or NH_4^+ can interact with liquid smoke species to artificially introduce N through adduction or carbonyl reactions. Methanol solvent can also potentially alter the functionalities of carbonyls, although it is less severe due to the lower abundances of D_3 and D_6 species. However, using isotopically labeled solvents or reagents in tandem with ultrahigh resolution MS and data science tools can also offer insight into complex sample composition. For example, based on the HDX studies, it is likely that there are significant numbers of enol, secondary N, and carboxylic acid species in mesquite liquid smoke. In addition, the artifact studies suggest there may be significant numbers of carbonyl compounds, although it is difficult to confirm these functionalities without more structural information.

6.1 Hydrogen-Deuterium Exchange Conclusions

Hydrogen deuterium exchange (HDX), a process where labile protons are exchanged for deuterium, occurs readily in mesquite liquid smoke with up to 3 labile hydrogens being exchanged in some cases. There is a difference between positive and negative ions, as well as between H^+ and Na^+ adducted species in positive mode. HDX is significantly detected for negative CHO compounds. Species with high numbers of HDX

may be rich in carboxyl functional groups and species with lower numbers of HDX may have fewer carboxyl groups and be more influenced by enols due to tautomerization. Similarly, Na^+ adducted HDX-capable species had solely CHO compositions, but were less significant since there were fewer Na^+ formulas and formulas capable of exchange. H^+ adducted positive ions favored N-containing compounds with less correlation from O and H/C values. These species tended to have lower numbers of exchange, implying secondary amines, amides, or aromatic N species. In all, it is necessary to use both ionization modes in order to see the full range of compounds capable of undergoing exchange.

Currently, the presence of HDX was investigated for only the most likely compounds, which involved omitting many molecular formulas from analysis that may have been chemically valid. Future investigations could involve an application of more decisive filters for formula assignment to increase the confidence in higher numbers of exchange and determining the difference between OH and COOH groups. Relative abundance, for example, could be used to filter molecular formulas further, as was found in the standard mix experiment, where several acidic organic compounds were found to have differences in their HDX abundance patterns. More information on HDX rates in individual compounds is needed to consider relative abundance.

6.2 NH₄OH Artifact Conclusions

The use of NH₄OH during the solid-phase extraction of complex organic mixtures results in significant reaction artifacts or adducts between CHO/CHNO ions in the sample and NH₃ from the NH₄OH. These interactions impact the analysis of samples obtained in both positive and negative mode ESI, specifically for liquid smoke samples, which are biomass burning surrogates. Analyses utilizing SPE sample prep will need to be concerned with interpreting the second elution data, as the use of NH₄OH significantly alters the sample composition by converting CHO compounds into CHNO compounds. Because this reaction can proceed through many different reaction routes, it is difficult to be certain if one reaction occurs preferentially over the other based on the available data. The molecular formula composition of the potential precursors is very similar in terms of the average DBE, number of O, and mass. Adduct formation precursors, however, can be distinguished from reaction artifact precursors based on aromaticity.

Although non-SPE samples were prepared with varying amounts of formic acid, the results for these samples were inconclusive for both ion polarities. The effect of acidity was considered in this work due to previous analysis, where preliminary evidence suggested that adding formic acid reduced the severity of the artifact formation. However, this was not observed here. In this regard, it is worth investigating other samples using similar parameters to further investigate this phenomenon. Additional concentrations of acid could also be used to determine when acid *does* significantly influence the artifact formation.

To improve the differentiation of the various possible reaction pathway precursors, future work could apply predictive separation methods such as principal component analysis to distinguish which parameters influence the possibility of one reaction occurring over another, especially in terms of ester hydrolysis and imine formation. Regardless, it is difficult to generate a detailed understanding of the true artifact reaction pathway due to the very complex nature of these mixtures. Since many of the observed reaction precursors exist naturally in complex mixtures and their abundances are so variable, the route for artifact production may vary from sample to sample. Future investigations could also be performed using $^{15}\text{NH}_4\text{OH}$ to study the formation and reactivity of different functional groups, such as imidazoles.

To obtain an improved understanding of the mechanism of NH_4OH artifacts, future analysis could be done on samples with different methods of ionization, such as atmospheric pressure chemical ionization or photoionization. These methods produce ions from more volatile species and may offer additional insight into whether these artifacts are truly dependent on reactions in solution or during the ionization.

6.3 Methanol Artifact Conclusions

Reaction artifacts in MeOH-solvated mesquite liquid smoke occur in detectable amounts, as indicated by the significant number of D-containing formulas that were traceable to reaction precursors using MeOH- d_3 . This occurs in both positive and negative ESI, indicating that the reaction artifact occurs in solution, although there is evidence suggesting the positive electrospray process accelerates the reactions. The

overall extent of artifact formation is relatively minor, as the total ion intensities were relatively low in comparison to the natural amounts of the reaction products in regular MeOH. In addition, reactions tended to occur with the more aliphatic and olefinic ions and was not deemed to be significant for aromatic species. This indicates that methanol reactions with carbonyls are less likely in aromatic species due to the increased electronic stability of the adjacent aromatic rings which impacts the carbonyl stability. Methanol reaction artifacts do not significantly impact the observed composition of mesquite liquid smoke. In turn, this implies that biomass burning compositions are not significantly affected by these reaction pathways, as they are rich in phenols and other aromatic species.

Although only the most likely reaction precursors were found, it is possible that MeOH may promote other reactions. Some D-containing formulas of significant intensity (*i.e.* more intense than the found precursor) were removed from this analysis, and they could be linked to such reactions. This would require a more thorough investigation with more potential reaction pathways and could be the focus of future work. More accurately separating the products of the ester and hemiacetal pathways could also be the focus of future study.

7 Reference List

- B&G Foods, I. (2019). Retrieved from <https://wrightsliquidsmoke.com/>
- Bateman, A. P., Walser, M. L., Desyaterik, Y., Laskin, J., Laskin, A., & Nizkorodov, S. A. (2008). The Effect of Solvent on the Analysis of Secondary Organic Aerosol Using Electrospray Ionization Mass Spectrometry. *Environmental Science & Technology*, 42(19), 7341-7346. <http://dx.doi.org/10.1021/es801226w>
- Baugh, J., Ray, W., & Black, F. (1987). *Motor vehicle emissions under reduced ambient-temperature idle-operating conditions (journal version)* (PB-89-118475/XAB; EPA-600/J-87/445 United States NTIS, PC A02/MF A01. GRA English). Retrieved from
- Brege, M., Paglione, M., Gilardoni, S., Decesari, S., Facchini, M. C., & Mazzoleni, L. R. (2018). Molecular insights on aging and aqueous-phase processing from ambient biomass burning emissions-influenced Po Valley fog and aerosol. *Atmos. Chem. Phys.*, 18(17), 13197-13214. <https://www.atmos-chem-phys.net/18/13197/2018/>
- Brege, M. C., S.; Mazzoleni, L. (2019). *Extreme molecular complexity in biomass burning atmospheric tar balls observed using complimentary soft ionization methods and ultrahigh resolution mass spectrometry*. In Prep.
- Brenton, A. G., & Godfrey, A. R. (2010). Accurate Mass Measurement: Terminology and Treatment of Data. *J Am Soc Mass Spectrom*, 21(11), 1821-1835. <http://www.sciencedirect.com/science/article/pii/S1044030510004022>
- Budisulistiorini, S. H., Riva, M., Williams, M., Chen, J., Itoh, M., Surratt, J. D., & Kuwata, M. (2017). Light-Absorbing Brown Carbon Aerosol Constituents from Combustion of Indonesian Peat and Biomass. *Environmental Science & Technology*, 51(8), 4415-4423. <https://doi.org/10.1021/acs.est.7b00397>
- Cech, N. B., & Enke, C. G. (2001). Effect of Affinity for Droplet Surfaces on the Fraction of Analyte Molecules Charged during Electrospray Droplet Fission. *Anal Chem*, 73(19), 4632-4639. <https://doi.org/10.1021/ac001267j>
- Cech, N. B., & Enke, C. G. (2001). Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrom Rev*, 20(6), 362-387.
- Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., et al. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet*, 389(10082), 1907-1918.
- De Haan, D. O., Hawkins, L. N., Kononenko, J. A., Turley, J. J., Corrigan, A. L., Tolbert, M. A., & Jimenez, J. L. (2011). Formation of Nitrogen-Containing Oligomers by Methylglyoxal and Amines in Simulated Evaporating Cloud Droplets.

- Environmental Science & Technology*, 45(3), 984-991.
<https://doi.org/10.1021/es102933x>
- Decesari, S., Facchini, M. C., Mircea, M., Cavalli, F., & Fuzzi, S. (2003). Solubility properties of surfactants in atmospheric aerosol and cloud/fog water samples. *108(D21)*.
<https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003JD003566>
- Ehrmann, B. M., Henriksen, T., & Cech, N. B. (2008). Relative importance of basicity in the gas phase and in solution for determining selectivity in electrospray ionization mass spectrometry. *J Am Soc Mass Spectrom*, 19(5), 719-728.
<https://doi.org/10.1016/j.jasms.2008.01.003>
- Farmer, D. K., Matsunaga, A., Docherty, K. S., Surratt, J. D., Seinfeld, J. H., Ziemann, P. J., & Jimenez, J. L. (2010). Response of an aerosol mass spectrometer to organonitrates and organosulfates and implications for atmospheric chemistry. *107(15)*, 6670-6675. <https://www.pnas.org/content/pnas/107/15/6670.full.pdf>
- Fenn, J., Mann, M., Meng, C., Wong, S., & Whitehouse, C. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *246(4926)*, 64-71.
<http://science.sciencemag.org/content/sci/246/4926/64.full.pdf>
- Flerus, R., Koch, B. P., Schmitt-Kopplin, P., Witt, M., & Kattner, G. (2011). Molecular level investigation of reactions between dissolved organic matter and extraction solvents using FT-ICR MS. *Marine Chemistry*, 124(1-4), 100-107.
<https://doi.org/10.1016/j.marchem.2010.12.006>
- Fry, J. L., Kiendler-Scharr, A., Rollins, A. W., Wooldridge, P. J., Brown, S. S., Fuchs, H., et al. (2009). Organic nitrate and secondary organic aerosol yield from NO₃ oxidation of β -pinene evaluated using a gas-phase kinetics/aerosol partitioning model. *Atmos. Chem. Phys.*, 9(4), 1431-1449.
<https://www.atmos-chem-phys.net/9/1431/2009/>
- Galloway, M. M., Chhabra, P. S., Chan, A. W. H., Surratt, J. D., Flagan, R. C., Seinfeld, J. H., & Keutsch, F. N. (2009). Glyoxal uptake on ammonium sulphate seed aerosol: reaction products and reversibility of uptake under dark and irradiated conditions. *Atmos. Chem. Phys.*, 9(10), 3331-3345. <https://www.atmos-chem-phys.net/9/3331/2009/>
- Gao, Y., Hall, W. A., & Johnston, M. V. (2010). Molecular Composition of Monoterpene Secondary Organic Aerosol at Low Mass Loading. *Environmental Science & Technology*, 44(20), 7897-7902. <https://doi.org/10.1021/es101861k>
- Gavard, R., Rossell, D., Spencer, S. E. F., & Barrow, M. P. (2017). Themis: Batch Preprocessing for Ultrahigh-Resolution Mass Spectra of Complex Mixtures. *Anal Chem*, 89(21), 11383-11390. <https://doi.org/10.1021/acs.analchem.7b02345>

- Hawkins, L. N., Welsh, H. G., & Alexander, M. V. (2018). Evidence for pyrazine-based chromophores in cloud water mimics containing methylglyoxal and ammonium sulfate. *Atmos. Chem. Phys.*, *18*(16), 12413-12431. <https://www.atmos-chem-phys.net/18/12413/2018/>
- Heald, C. L., Kroll, J. H., Jimenez, J. L., Docherty, K. S., DeCarlo, P. F., Aiken, A. C., et al. (2010). A simplified description of the evolution of organic aerosol composition in the atmosphere. *37*(8). <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010GL042737>
- Henriksen, T., Juhler, R. K., Svensmark, B., & Cech, N. B. (2005). The relative influences of acidity and polarity on responsiveness of small organic molecules to analysis with negative ion electrospray ionization mass spectrometry (ESI-MS). *J Am Soc Mass Spectrom*, *16*(4), 446-455.
- Herzsprung, P., Hertkorn, N., von Tumpling, W., Harir, M., Friese, K., & Schmitt-Kopplin, P. (2014). Understanding molecular formula assignment of Fourier transform ion cyclotron resonance mass spectrometry data of natural organic matter from a chemical point of view. *Anal Bioanal Chem*, *406*(30), 7977-7987.
- Hoffmann, E. d. (2005). Mass Spectrometry. In *Kirk-Othmer Encyclopedia of Chemical Technology*.
- Hossain, B. M., & Konermann, L. (2006). Pulsed hydrogen/deuterium exchange MS/MS for studying the relationship between noncovalent protein complexes in solution and in the gas phase after electrospray ionization. *Anal Chem*, *78*(5), 1613-1619.
- IPCC. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Retrieved from
- Isaacman, G., Wilson, K. R., Chan, A. W., Worton, D. R., Kimmel, J. R., Nah, T., et al. (2012). Improved resolution of hydrocarbon structures and constitutional isomers in complex mixtures using gas chromatography-vacuum ultraviolet-mass spectrometry. *Anal Chem*, *84*(5), 2335-2342.
- Kendrick, E. (1963). A Mass Scale Based on CH₂ = 14.0000 for High Resolution Mass Spectrometry of Organic Compounds. *Anal Chem*, *35*(13), 2146-2154. <https://doi.org/10.1021/ac60206a048>
- Kind, T., & Fiehn, O. (2007). Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC bioinformatics*, *8*, 105-105. <https://www.ncbi.nlm.nih.gov/pubmed/17389044>
<https://www.ncbi.nlm.nih.gov/pmc/PMC1851972/>

- Koch, B. P., Dittmar, T., Witt, M., & Kattner, G. (2007). Fundamentals of Molecular Formula Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter. *Anal Chem*, 79(4), 1758-1763. <https://doi.org/10.1021/ac061949s>
- Kostyukevich, Y., Acter, T., Zherebker, A., Ahmed, A., Kim, S., & Nikolaev, E. (2018). Hydrogen/deuterium exchange in mass spectrometry. *Mass Spectrom Rev*, 37(6), 811-853. <https://doi.org/10.1002/mas.21565>
- Kostyukevich, Y., Kononikhin, A., Popov, I., Kharybin, O., Perminova, I., Konstantinov, A., & Nikolaev, E. (2013). Enumeration of Labile Hydrogens in Natural Organic Matter by Use of Hydrogen/Deuterium Exchange Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal Chem*, 85(22), 11007-11013. <https://doi.org/10.1021/ac402609x>
- Kostyukevich, Y., Kononikhin, A., Popov, I., & Nikolaev, E. (2013). Simple Atmospheric Hydrogen/Deuterium Exchange Method for Enumeration of Labile Hydrogens by Electrospray Ionization Mass Spectrometry. *Anal Chem*, 85(11), 5330-5334. <https://doi.org/10.1021/ac4006606>
- Kozhinov, A. N., Zhurov, K. O., & Tsybin, Y. O. (2013). Iterative Method for Mass Spectra Recalibration via Empirical Estimation of the Mass Calibration Function for Fourier Transform Mass Spectrometry-Based Petroleomics. *Anal Chem*, 85(13), 6437-6445. <https://doi.org/10.1021/ac400972y>
- Krueve, A., Kaupmees, K., Liigand, J., Oss, M., & Leito, I. (2013). Sodium adduct formation efficiency in ESI source. 48(6), 695-702. <https://onlinelibrary.wiley.com/doi/abs/10.1002/jms.3218>
- Krueve, A., & Kaupmees, K. J. J. o. T. A. S. f. M. S. (2017). Adduct Formation in ESI/MS by Mobile Phase Additives. 28(5), 887-894. journal article. <https://doi.org/10.1007/s13361-017-1626-y>
- Kujawinski, E. B. (2002). Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (ESI FT-ICR MS): Characterization of Complex Environmental Mixtures. *Environmental Forensics*, 3(3-4), 207-216. <https://www.tandfonline.com/doi/abs/10.1080/713848382>
- Laskin, A., Smith, J. S., & Laskin, J. (2009). Molecular Characterization of Nitrogen-Containing Organic Compounds in Biomass Burning Aerosols Using High-Resolution Mass Spectrometry. *Environmental Science & Technology*, 43(10), 3764-3771. <https://doi.org/10.1021/es803456n>
- Liigand, J., Krueve, A., Leito, I., Girod, M., & Antoine, R. (2014). *Effect of Mobile Phase on Electrospray Ionization Efficiency* (Vol. 25).
- Lin, P., Fleming, L. T., Nizkorodov, S. A., Laskin, J., & Laskin, A. (2018). Comprehensive Molecular Characterization of Atmospheric Brown Carbon by High Resolution Mass Spectrometry with Electrospray and Atmospheric Pressure

- Photoionization. *Anal Chem*, 90(21), 12493-12502.
<https://doi.org/10.1021/acs.analchem.8b02177>
- Lin, P., Laskin, J., Nizkorodov, S. A., & Laskin, A. (2015). Revealing Brown Carbon Chromophores Produced in Reactions of Methylglyoxal with Ammonium Sulfate. *Environmental Science & Technology*, 49(24), 14257-14266.
<https://doi.org/10.1021/acs.est.5b03608>
- Lin, P., Yu, J. Z., Engling, G., & Kalberer, M. (2012). Organosulfates in Humic-like Substance Fraction Isolated from Aerosols at Seven Locations in East Asia: A Study by Ultra-High-Resolution Mass Spectrometry. *Environmental Science & Technology*, 46(24), 13118-13127. <https://doi.org/10.1021/es303570v>
- Makarov, A., Denisov, E., Lange, O., & Horning, S. J. J. o. T. A. S. f. M. S. (2006). Dynamic range of mass accuracy in LTQ orbitrap hybrid mass spectrometer. 17(7), 977-982. journal article. <https://doi.org/10.1016/j.iasms.2006.03.006>
- Mazzoleni, L., Saranjampour, P., Dalbec, M., Samburova, V., A Gannet, D., Zielinska, B., et al. (2012). *Identification of water-soluble organic carbon in non-urban aerosols using ultrahigh-resolution FT-ICR mass spectrometry: Organic anions* (Vol. 9).
- Mazzoleni, L. R., Ehrmann, B. M., Shen, X., Marshall, A. G., & Collett, J. L. (2010). Water-Soluble Atmospheric Organic Matter in Fog: Exact Masses and Chemical Formula Identification by Ultrahigh-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Environmental Science & Technology*, 44(10), 3690-3697. <https://doi.org/10.1021/es903409k>
- McIntyre, C., McRae, C., Jardine, D., & Batts, B. D. (2002). Self-esterification of fulvic acid model compounds in methanolic solvents as observed by electrospray ionization mass spectrometry. *Rapid Commun Mass Spectrom*, 16(8), 785-789.
- Michalski, A., Damoc, E., Lange, O., Denisov, E., Nolting, D., Muller, M., et al. (2012). Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Mol Cell Proteomics*, 11(3), O111.013698.
- Montazeri, N., Oliveira, A. C., Himelbloom, B. H., Leigh, M. B., & Crapo, C. A. (2013). Chemical characterization of commercial liquid smoke products. *Food science & nutrition*, 1(1), 102-115. <https://www.ncbi.nlm.nih.gov/pubmed/24804019>
<https://www.ncbi.nlm.nih.gov/pmc/PMC3951573/>
- Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Chhabra, P. S., Seinfeld, J. H., & Worsnop, D. R. (2011). Changes in organic aerosol composition with aging inferred from aerosol mass spectra. *Atmos. Chem. Phys.*, 11(13), 6465-6474.
<https://www.atmos-chem-phys.net/11/6465/2011/>

- Nizkorodov, S. A., Laskin, J., & Laskin, A. (2011). Molecular chemistry of organic aerosols through the application of high resolution mass spectrometry. *Physical Chemistry Chemical Physics*, 13(9), 3612-3629. 10.1039/C0CP02032J. <http://dx.doi.org/10.1039/C0CP02032J>
- Novotny, N. R., Capley, E. N., & Stenson, A. C. (2014). Fact or artifact: the representativeness of ESI-MS for complex natural organic mixtures. *Journal of Mass Spectrometry*, 49(4), 316-326. <https://doi.org/10.1002/jms.3345>
- Oganesyan, I., Lento, C., & Wilson, D. J. (2018). Contemporary hydrogen deuterium exchange mass spectrometry. *Methods*, 144, 27-42. <http://www.sciencedirect.com/science/article/pii/S1046202317304528>
- Ohno, T., Ohno, P. E. J. A., & Chemistry, B. (2013). Influence of heteroatom pre-selection on the molecular formula assignment of soil organic matter components determined by ultrahigh resolution mass spectrometry. *405*(10), 3299-3306. journal article. <https://doi.org/10.1007/s00216-013-6734-3>
- Perdue, E. M. (2013). *Standard and Reference Samples of Humic Acids, Fulvic Acids, and Natural Organic Matter from the Suwannee River, Georgia: Thirty Years of Isolation and Characterization*, Dordrecht.
- Purcell, J. M., Hendrickson, C. L., Rodgers, R. P., & Marshall, A. G. (2007). Atmospheric Pressure Photoionization Proton Transfer for Complex Organic Mixtures Investigated by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J Am Soc Mass Spectrom*, 18(9), 1682-1689. <http://www.sciencedirect.com/science/article/pii/S1044030507005557>
- Ramanathan, R., Cao, K., Cavalieri, E., & Gross, M. L. (1998). Mass spectrometric methods for distinguishing structural isomers of glutathione conjugates of estrone and estradiol. *J Am Soc Mass Spectrom*, 9(6), 612-619.
- Reemtsma, T., These, A., Venkatachari, P., Xia, X., Hopke, P. K., Springer, A., & Linscheid, M. (2006). Identification of Fulvic Acids and Sulfated and Nitrated Analogues in Atmospheric Aerosol by Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal Chem*, 78(24), 8299-8304. <https://doi.org/10.1021/ac061320p>
- Remer, L. A., & Kaufman, Y. J. (1998). Dynamic aerosol model: Urban/industrial aerosol. *103*(D12), 13859-13871. <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JD00994>
- Reusch, W. (1999). *Virtual Textbook of Organic Chemistry*. In.
- Schum, S.K., L.R. Mazzoleni, L.E. Brown (2018). MFAssignR: An R package for Data Preparation and Molecular Formula Assignment, R package version 0.0.2, <http://github.com/skschum/MFAssignR>, doi: 10.5281/zenodo.1471471.

- Schum, S. K. (2019). *Molecular characterization of free-tropospheric organic aerosol and the development of computational tools for molecular formula assignments* (Unpublished Dissertation), Michigan Technological University, Houghton, MI, USA.
- Shiraiwa, M., Ueda, K., Pozzer, A., Lammel, G., Kampf, C. J., Fushimi, A., et al. (2017). Aerosol Health Effects from Molecular to Global Scales. *Environmental Science & Technology*, 51(23), 13545-13567. <https://doi.org/10.1021/acs.est.7b04417>
- Sleighter, R. L., McKee, G. A., Liu, Z., & Hatcher, P. G. (2008). Naturally present fatty acids as internal calibrants for Fourier transform mass spectra of dissolved organic matter. 6(6), 246-253. <https://aslopubs.onlinelibrary.wiley.com/doi/abs/10.4319/lom.2008.6.246>
- Smith, J. S., Laskin, A., & Laskin, J. (2009). Molecular Characterization of Biomass Burning Aerosols Using High-Resolution Mass Spectrometry. *Anal Chem*, 81(4), 1512-1521. <https://doi.org/10.1021/ac8020664>
- Smith, R. D., Loo, J. A., Edmonds, C. G., Barinaga, C. J., & Udseth, H. R. (1990). New developments in biochemical mass spectrometry: electrospray ionization. *Anal Chem*, 62(9), 882-899. <https://doi.org/10.1021/ac00208a002>
- Stanz, D. (2015). Composer64 v1.5.0: Sierra Analytics, Inc.
- Stenson, A. C., Marshall, A. G., & Cooper, W. T. (2003). Exact Masses and Chemical Formulas of Individual Suwannee River Fulvic Acids from Ultrahigh Resolution Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectra. *Anal Chem*, 75(6), 1275-1284. <https://doi.org/10.1021/ac026106p>
- Stenson, A. C., Ruddy, B. M., & Bythell, B. J. (2014). Ion molecule reaction H/D exchange as a probe for isomeric fractionation in chromatographically separated natural organic matter. 360. <https://doi.org/10.1016/j.ijms.2013.12.026>
- Sun, Y. L., Zhang, Q., Anastasio, C., & Sun, J. (2010). Insights into secondary organic aerosol formed via aqueous-phase reactions of phenolic compounds based on high resolution mass spectrometry. *Atmos. Chem. Phys.*, 10(10), 4809-4822. <https://www.atmos-chem-phys.net/10/4809/2010/>
- Taflin, D. C., Ward, T. L., & Davis, E. J. (1989). Electrified droplet fission and the Rayleigh limit. *Langmuir*, 5(2), 376-384. <https://doi.org/10.1021/la00086a016>
- Teich, M., van Pinxteren, D., Kecorius, S., Wang, Z., & Herrmann, H. (2016). First Quantification of Imidazoles in Ambient Aerosol Particles: Potential Photosensitizers, Brown Carbon Constituents, and Hazardous Components. *Environmental Science & Technology*, 50(3), 1166-1173. <https://doi.org/10.1021/acs.est.5b05474>

- Tolocka, M. P., Jang, M., Ginter, J. M., Cox, F. J., Kamens, R. M., & Johnston, M. V. (2004). Formation of Oligomers in Secondary Organic Aerosol. *Environmental Science & Technology*, 38(5), 1428-1434. <https://doi.org/10.1021/es035030r>
- Westerling, A. L., Hidalgo, H. G., Cayan, D. R., & Swetnam, T. W. (2006). Warming and Earlier Spring Increase Western U.S. Forest Wildfire Activity. *313(5789)*, 940-943. <http://science.sciencemag.org/content/sci/313/5789/940.full.pdf>
- Wickham, H. (Producer). (2016). ggplot2: Elegant Graphics for Data Analysis.
- Wu, J., & McAllister, H. (2003). Exact mass measurement on an electrospray ionization time-of-flight mass spectrometer: error distribution and selective averaging. *38(10)*, 1043-1053. <https://onlinelibrary.wiley.com/doi/abs/10.1002/jms.516>
- Yan, X., & Maier, C. S. (2009). Hydrogen/Deuterium Exchange Mass Spectrometry. In M. S. Lipton & L. Paša-Tolic (Eds.), *Mass Spectrometry of Proteins and Peptides: Methods and Protocols* (pp. 255-271). Totowa, NJ: Humana Press.
- Yasmeen, F., Szmigielski, R., Vermeylen, R., Gomez-Gonzalez, Y., Surratt, J. D., Chan, A. W., et al. (2011). Mass spectrometric characterization of isomeric terpenoic acids from the oxidation of alpha-pinene, beta-pinene, d-limonene, and Delta3-carene in fine forest aerosol. *J Mass Spectrom*, 46(4), 425-442.
- Yasmeen, F., Vermeylen, R., Szmigielski, R., Iinuma, Y., Böge, O., Herrmann, H., et al. (2010). Terpenylic acid and related compounds: precursors for dimers in secondary organic aerosol from the ozonolysis of α- and β-pinene. *Atmos. Chem. Phys.*, 10(19), 9383-9392. <https://www.atmos-chem-phys.net/10/9383/2010/>
- Zark, M., Christoffers, J., & Dittmar, T. (2017). Molecular properties of deep-sea dissolved organic matter are predictable by the central limit theorem: Evidence from tandem FT-ICR-MS. *Marine Chemistry*, 191, 9-15. <http://www.sciencedirect.com/science/article/pii/S0304420316301189>
- Zhang, J., & Smith, K. R. (1999). Emissions of Carbonyl Compounds from Various Cookstoves in China. *Environmental Science & Technology*, 33(14), 2311-2320. <https://doi.org/10.1021/es9812406>
- Zhao, J., Ortega, J., Chen, M., McMurry, P. H., & Smith, J. N. (2013). Dependence of particle nucleation and growth on high-molecular-weight gas-phase products during ozonolysis of α -pinene. *Atmos. Chem. Phys.*, 13(15), 7631-7644. <https://www.atmos-chem-phys.net/13/7631/2013/>
- Zhou, S., Prebyl, B. S., & Cook, K. D. (2002). Profiling pH Changes in the Electrospray Plume. *Anal Chem*, 74(19), 4885-4888. <https://doi.org/10.1021/ac025960d>