**Michigan
Technological
University**

Michigan Technological University
Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's Reports

2019

EXTREME MOLECULAR DIVERSITY IN BIOMASS BURNING ATMOSPHERIC ORGANIC AEROSOL OBSERVED THROUGH ULTRAHIGH RESOLUTION MASS SPECTROMETRY

Matthew Brege

Michigan Technological University, mabrege@mtu.edu

Copyright 2019 Matthew Brege

Recommended Citation

Brege, Matthew, "EXTREME MOLECULAR DIVERSITY IN BIOMASS BURNING ATMOSPHERIC ORGANIC AEROSOL OBSERVED THROUGH ULTRAHIGH RESOLUTION MASS SPECTROMETRY", Open Access Dissertation, Michigan Technological University, 2019.
<https://doi.org/10.37099/mtu.dc.etdr/927>

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

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

EXTREME MOLECULAR DIVERSITY IN BIOMASS BURNING ATMOSPHERIC
ORGANIC AEROSOL OBSERVED THROUGH ULTRAHIGH RESOLUTION MASS
SPECTROMETRY

By

Matthew A. Brege

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Chemistry

MICHIGAN TECHNOLOGICAL UNIVERSITY

2019

© 2019 Matthew A. Brege

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Chemistry.

Department of Chemistry

Dissertation Advisor: *Lynn R. Mazzoleni*

Committee Member: *Will H. Cantrell*

Committee Member: *Sarah A. Green*

Committee Member: *Marina Tanasova*

Department Chair: *John A. Jaszczak*

Table of Contents

Preface.....	vi
Acknowledgements.....	vii
Abstract.....	ix
1 Introduction.....	1
1.1 Introduction to atmospheric aerosol.....	1
1.1.1 Atmospheric aerosol and climate.....	1
1.1.2 Atmospheric aerosol and human health.....	3
1.1.3 Sources and composition of atmospheric aerosol.....	3
1.1.4 Aerosol sampling.....	7
1.2 Introduction to mass spectrometry.....	7
1.2.1 Molecular mass.....	7
1.2.2 Ionization of analyte molecules.....	8
1.2.3 Ultrahigh resolution mass spectrometry.....	10
1.2.3.1 Fourier transform-ion cyclotron resonance instruments.....	12
1.2.3.2 Orbitrap instruments.....	12
1.2.4 Sample preparation.....	13
1.2.5 Post processing and data analysis.....	13
1.2.5.1 Molecular formula assignment.....	14
1.2.5.2 Molecular formula validation.....	16
2 Molecular insights on aging and aqueous phase processing from ambient biomass burning emissions-influenced Po Valley fog and aerosol.....	18
2.1 Abstract.....	18
2.2 Introduction.....	19
2.3 Methods.....	21
2.3.1 Sample collection and chemical analysis.....	21
2.3.2 ¹ H-NMR analysis.....	22
2.3.3 HR-ToF-AMS analysis.....	22
2.3.4 Ultrahigh resolution FT-ICR MS analysis.....	23
2.4 Results and discussion.....	24
2.4.1 Selection of aerosol and fog water samples.....	24
2.4.2 ¹ H-NMR composition.....	26
2.4.3 Ultrahigh resolution FT-ICR MS composition.....	30
2.4.3.1 Overview of the Po Valley ambient fog and aerosol compositions.....	30
2.4.3.2 Molecular trends for ambient fog and aerosol compositions.....	38

2.4.3.3	Comparison of the fresh and aged biomass burning influenced fog compositions.....	41
2.4.3.4	Comparison of the fresh and aged biomass burning influenced aerosol compositions.....	45
2.5	Summary and implications.....	47
2.6	Associated content.....	49
2.6.1	Data availability.....	49
2.6.2	Author contributions.....	49
2.6.3	Acknowledgments.....	49
2.7	Supplementary information: chapter 2.....	49
2.7.1	FT-ICR MS data processing and molecular formula assignment review.....	49
2.7.2	Ultrahigh resolution FT-ICR MS results.....	50
3	Molecular trends observed in sequential samples of biomass burning-influenced aerosol and fog water in the Po Valley analyzed by ultrahigh resolution Orbitrap Elite mass spectrometry.....	52
3.1	Abstract.....	52
3.2	Introduction.....	53
3.3	Methods.....	55
3.4	Results and discussion.....	59
3.4.1	Comparing day to night samples.....	79
3.4.2	Comparing day to fog samples.....	79
3.4.3	Comparing night to fog samples.....	80
3.4.4	Pyrazine chromophores.....	80
3.5	Conclusions.....	84
3.6	Associated content.....	89
3.6.1	Data availability.....	89
3.6.2	Acknowledgements.....	89
4	Extreme molecular complexity resulting in a continuum of carbonaceous species in biomass burning tar balls.....	90
4.1	Abstract.....	90
4.2	Introduction.....	90
4.3	Methods.....	93
4.4	Results and Discussion.....	101
4.4.1	Molecular composition of tar balls.....	101
4.4.2	Volatility of tar ball molecular species.....	110
4.4.3	Health and atmospheric impacts.....	114

4.5	Associated content.....	119
4.5.1	Data availability	119
4.5.2	Acknowledgements.....	119
4.6	Supplementary information: chapter 4.....	119
4.6.1	Use of MFAssignR for molecular formula assignment	119
4.6.1.1	Signal-to-noise estimation	120
4.6.1.2	Spectra recalibration	120
4.6.1.3	Molecular formula assignment	122
4.6.1.4	Data quality assurance and molecular formula validation..	124
4.6.1.5	Blank subtraction	126
4.6.2	Estimation of chemical and physical properties	126
4.6.2.1	Double bond equivalents.....	126
4.6.2.2	Aromaticity index	126
4.6.2.3	Average oxidation state of carbon	127
4.6.2.4	Volatility estimation.....	127
4.6.2.5	Dry and relative humidity dependent glass transition temperatures.....	128
4.6.3	Molecular formulas observed in multiple ionization methods	129
5	Conclusion	133
5.1	Summary	133
5.2	Recommendations for further research	134
6	Appendix A: ancillary tables and figures.....	137
6.1	Ancillary tables.....	137
6.2	Ancillary figures.....	143
7	Appendix B: copyright documentation.....	156
8	References.....	157

Preface

Chapter 2 was published in the journal '*Atmospheric Chemistry and Physics*' (Brege et al., 2018) and is republished in its entirety here with permission from the lead author and the publisher. The sample collection was performed in collaboration with Dr. Marco Paglione, Dr. Stefania Gilardoni, Dr. Stefano Decesari and Dr. Maria Cristina Facchini; the proton nuclear magnetic resonance, the high resolution time of flight aerosol mass spectrometry and the positive matrix factorization analyses were performed by these collaborators. The Fourier transform ion cyclotron resonance mass spectrometry data collection and analysis were performed by myself and Dr. Lynn Mazzoleni. The manuscript was prepared by myself with significant contributions from all coauthors.

Chapter 3 was prepared with the intent to be submitted for publication in the spring of 2020 and is considered a work in progress. The sample collection was performed in collaboration with Dr. Marco Paglione, Dr. Stefania Gilardoni, Dr. Stefano Decesari and Dr. Maria Cristina Facchini. The Orbitrap Elite mass spectrometry data collection and analysis was performed by myself. The manuscript was prepared by myself with significant contributions from all coauthors.

Chapter 4 was prepared with the intent to be submitted to the journal '*Environmental Science & Technology*' in the fall of 2019 and is considered a work in progress. The sample collection, scanning electron microscopy analysis and evaporation kinetics experiments were performed in collaboration with Dr. Swarup China and Dr. Alla Zelenyuk-Imre. The Orbitrap Elite mass spectrometry data collection and analysis was performed by myself and Dr. Simeon Schum. The manuscript was prepared by myself with significant contributions from all coauthors.

Acknowledgements

My graduate studies have been an extremely humbling experience. None of this work was completed in a vacuum, and I would not have made it this far without the support and contributions from my mentors, colleagues and friends. My successes are the product of those people around me who helped when I asked them to, and more importantly, pushed me when I needed to be.

My advisor Dr. Lynn Mazzoleni has been instrumental to my success in research. She is an extremely motivated and brilliant scientist and I am thankful for the opportunity to work with her. We have shared successes and sorrows, and she has always motivated me to do my best and be just a little bit better. I am confident in the skills she has taught me, and I feel prepared for future endeavors because of her mentorship.

I would like to thank Dr. Claudio Mazzoleni and his research group for their kindness and intermittent wisdom.

I am grateful to my past and present research group members, Dr. Katja Dzepina, Dr. Elena Kirillova, Dr. Maryam Khaksari, Dr. Simeon Schum, Dr. Yunzhu Zhao, Tyler Leverton, Thusitha Divisekara, Amna Ijaz, DM Ashraf Habib, Trevor Curtis, Megan Dalbec, Senait Gebreeyesus, Alexandra Maday and Tyler Sawall, for their assistance and camaraderie. I would also like to thank all other previous Mazzoleni group members for laying the foundation for my research.

My collaborators Dr. Maria Cristina Facchini, Dr. Marco Paglione, Dr. Stefania Gilardoni, and Dr. Stefano Decesari, provided the essential samples for my research, have offered critical insights and were extremely hospitable during my first experience with international travel. Likewise, my collaborators Dr. Swarup China and Dr. Alla Zelenyuk have also provided samples for analysis and offered integral guidance. It has been a privilege to work with all of them.

I thank the staff of the Chemistry department at Michigan Tech and the previous and current chairs, Dr. Cary Chabalowski and Dr. John Jaszczak, for their support and years of funding. I would especially like to thank my teaching instructors Lorri Reilly and Aparna Pandey for helping me to realize my capabilities as a teacher.

I am grateful to my committee members Dr. Will Cantrell, Dr. Sarah Green, and Dr. Marina Tanasova for their careful review of this dissertation to ensure the results are presented clearly and the conclusions are sound.

I would like to thank the NASA Earth and Space Sciences Fellowship for three years of financial support, so I was able to prioritize my research above all other duties.

My parents have been a constant source of inspiration in my life and I would never have progressed this far without their constant support. I love them with all my heart.

Finally, my wife Erica can never be thanked enough.

Abstract

Ambient atmospheric aerosol is ubiquitous in the atmosphere, originating from a variety of natural and man-made sources. These microscopic particles have profound impacts on the global climate system as well as human health. The organic fraction of atmospheric aerosol is an extremely complex mixture which is not yet fully characterized. These unknown organic aerosol species contribute to the uncertainty in the effect of aerosol on climate and uncertainty in overall ambient aerosol toxicity. Light absorbing organic aerosol can interact with incoming solar radiation and contribute to atmospheric heating; however, the source apportionment and overall fate of these absorbing organic aerosol species are not fully understood. The burning of woody and vegetative materials (biomass) is expected to be one source, while secondary chemical reactions in aqueous phase aerosol and liquid water droplets are another. In this work, we have analyzed ambient samples from the Po Valley (Italy) and Pacific Northwest (USA) influenced by biomass burning. Using ultrahigh resolution mass spectrometry and subsequent molecular formula assignment, we observe an extreme level of molecular complexity in atmospheric aerosol. We make several key observations regarding both biomass burning organic aerosol and aqueous phase processing based on the molecular details and the observed elemental trends in the assigned formulas. We estimate oxidation levels, heteroatom functionalization, aromatic character, volatility and glass transition temperature based on reliable molecular formula assignments. Overall, this work describes a level of complexity in organic aerosol much greater than previously indicated. We suspect that any one analytical technique is likely to miss certain aspects of this mixture, and that a variety of analytical methods must be employed to fully characterize and resolve the complex mixture in atmospheric organic aerosol.

1 Introduction

1.1 Introduction to atmospheric aerosol

Invisible to the naked eye, microscopic aerosol particles suspended in Earth's atmosphere have a profound impact on visibility, human health and global climate (WHO 2016, Moses and Akinyemi 2017, IPCC 2018). While there are anthropogenic sources of atmospheric aerosol, the majority are formed through natural and biogenic processes (Goldstein and Galbally 2007, Shrivastava et al. 2017). Particles can have a large variety of sizes, chemical compositions and morphological shapes which each play distinct roles in deciphering their origins, lifetimes and physical properties (Poschl 2005, Seinfeld and Pandis 2016). Given the proper humidity conditions aerosol can lead to the formation of haze, fog and clouds (Ciobanu et al. 2010, Gilardoni et al. 2014, Han et al. 2015), which can limit atmospheric visibility and affect satellite and radio signals as well as restrict air and ground travel. Precipitation from clouds is highly dependent on the formation of ice crystals, in which aerosol particles play an important role (Khain et al. 2000). While many of these interactions between aerosol and clouds are the products of naturally occurring processes, the presence of light absorbing aerosol plays a more concerning role in the changing global climate (IPCC 2018).

1.1.1 Atmospheric aerosol and climate

The global climate exists in a delicate balance of incoming and outgoing radiation (Trenberth et al. 2009). The Earth is warmed by incoming solar radiation at the surface, and then subsequently emits black body radiation back to outer space (Salby 2012). The energy conversion involved in generating heat causes the radiation emitted by the Earth to be a much longer wavelength than incoming solar radiation, thus referred to as *short-wave* and *long-wave* respectively (Trenberth et al. 2009, Salby 2012). A majority of short-wave solar radiation passes through the atmosphere unabsorbed until reaching the surface, whereas some of the outgoing long-wave radiation emitted by the Earth is reabsorbed in the atmosphere and can lead to additional heating (Salby 2012, Anderson et al. 2016). This “greenhouse effect” has been thoroughly studied for well mixed atmospheric gasses like carbon dioxide (CO₂), where increasing concentrations of these gasses which absorb long-wave radiation have been linked to atmospheric warming (Anderson et al. 2016, IPCC 2018). Regarding atmospheric aerosol particles, interactions with incoming short-wave solar radiation can lead to cooling by light scattering or warming by light absorption (Poschl 2005, Saleh et al. 2018). Whether an aerosol particle will scatter or absorb light depends on the chemical composition of the individual particle (Poschl 2005, Wang et al. 2013), while the size, shape and mixing state of the internal components also contribute to the magnitude of light absorption (Moosmuller et al. 2009, Shiraiwa et al. 2010).

Aerosol interactions with solar radiation either by absorption or scattering are known as the *direct effect of aerosol* (Kanakidou et al. 2005, Moosmuller et al. 2009, IPCC 2018). More complex interactions with clouds can lead to additional heating or cooling effects

and are known as the *indirect effect of aerosol* (Kanakidou et al. 2005, Chandrakar et al. 2016, IPCC 2018). Briefly, hygroscopic aerosol can lead to the formation of cloud droplets which can cause additional light scattering and cooling effects. Relatively higher number concentrations of these aerosol can also lead to the formation of more cloud droplets smaller in size which can lead to increased reflectivity and increased cloud lifetime (Twomey 1977, Feingold et al. 2003, Lohmann 2006, Rosenfeld 2006). Concurrently, the presence of light absorbing aerosol species in cloud droplets may lead to increased temperatures and faster evaporation of the individual droplets leading to the dissipation of the cloud (Ackerman et al. 2000).

Typically, inorganic species are associated with cooling by light scattering, and carbonaceous species are associated with warming by light absorption, but these are broad generalizations. The distinction between “carbonaceous” and “organic” chemical species in atmospheric aerosol is made due to the presence of “black carbon”. With large macromolecular condensed ring chemical structures, black carbon resembles elemental or graphitic carbon (Gelencsér 2005) and the three terms are not intended to be used interchangeably. While in a strictly technical sense black carbon could be considered an organic chemical species as it is almost entirely composed of carbon atoms, in atmospheric science the term “organic aerosol” typically refers to molecular species with significantly smaller chemical structures and with a variety of carbon oxidation states. Black carbon and organic carbon are further described as refractory and non-refractory carbon respectively. Black carbon can absorb short-wave radiation nearly universally across the solar spectrum, thus it has the potential for atmospheric warming much greater than other sources (Andreae 2001, Jacobson 2002, Saleh et al. 2018). Those organic aerosol species which can also absorb solar radiation are designated as “brown carbon” and typically absorb in blue-violet and ultra violet wavelengths (Andreae and Gelencser 2006, Sun et al. 2007, Saleh et al. 2018). The absorption of brown carbon has been described to be from individual light absorbing molecules (Samburova et al. 2016, Hawkins et al. 2018, Lin et al. 2018) and also from charge transfer interactions between molecules (Phillips and Smith 2014, 2015). Deposition of these absorbing aerosol species can also lead to a decrease in the reflectivity of surfaces such as snow (Wang et al. 2013).

The source apportionment and environmental fate of brown carbon species in the atmosphere are still somewhat uncertain. The effect of aerosol-cloud interactions on climate are likewise unclear. These factors contribute to the relatively large level of uncertainty in the magnitude of the effect of aerosol on the radiation balance of the Earth (IPCC 2018). The unknown sources, lifetime and fate of organic species in atmospheric aerosol and their interactions with clouds largely contribute to the error in these estimations (Kanakidou et al. 2005, Shrivastava et al. 2017). It has been proposed that the mitigation of anthropogenic light absorbing aerosol would have a more rapid impact on rising global temperatures than the reduction of anthropogenic CO₂ (Jacobson 2002).

1.1.2 Atmospheric aerosol and human health

In addition to adverse climate effects, atmospheric aerosol can affect the health and livelihood of human beings. In 2016, The World Health Organization classified air pollution as the greatest environmental risk to public health, contributing directly to 1:9 deaths in 2012 (WHO 2016). Exposure to particulate matter of less than 2.5 micrometers in diameter (PM_{2.5}) has dangerous health effects even in low concentrations (Schwartz et al. 2002, Bennett et al. 2019), being linked to pulmonary and cardiovascular disease (WHO 2016). Unfortunately, the annual mean exposure limit for exposure to PM_{2.5} (10 µg m⁻³) and the 24-hour exposure limit (25 µg m⁻³) advised by the WHO are exceeded in many densely populated areas of the world (Ye et al. 2003, Sharma and Maloo 2005, Bennett et al. 2019).

Part of the overall toxicity of atmospheric aerosol can be linked to some organic chemical components. Reactive oxygen species, such as hydroxy radical (OH•), superoxide anion (O₂•⁻), and hydrogen peroxide (H₂O₂) are associated with oxidative stress in living cells and can cause multiple health effects (Lin and Yu 2011, Poschl and Shiraiwa 2015). The in vivo catalytic formation of these reactive oxygen species has been associated with certain fractions of ambient aerosol, in particular for Humic-like substances (Lin and Yu 2011), cooking aerosol and biomass burning emissions (Verma et al. 2015) and water soluble transition metals (See et al. 2007). Polycyclic aromatic hydrocarbons (PAHs) have been associated with reactive oxygen species in living cells (Yin et al. 2007, Valavanidis et al. 2010) in addition to being carcinogenic (Samburova et al. 2017). Biomass burning emissions include PAHs (Samburova et al. 2016, Alves et al. 2017, Lin et al. 2018) along with many other phenolic and aromatic chemical species (Mazzoleni et al. 2007, Kitanovski et al. 2012) that are likely related to the generation of reactive oxygen species (Verma et al. 2015). Other expected emissions include nitro-phenolic species which are known to be mutagenic and cytotoxic (Shimizu and Yano 1986, Harrison et al. 2005).

Forest fires in western North America are expected to increase in severity with rising global temperatures (Westerling et al. 2006, Barbero et al. 2015). The biomass burning emissions from these fires include both light absorbing black and brown carbon (Reid et al. 2005, Andreae and Gelencser 2006), which could further contribute to rising global temperatures in a positive feedback loop. Concurrently, it is possible that increasing global temperatures may contribute to increased aerosol toxicity from the same biomass burning emissions.

1.1.3 Sources and composition of atmospheric aerosol

Aerosol exist in size ranges from approximately 10 nm to 100 µm and can come from a wide variety of sources (Poschl and Shiraiwa 2015). Inorganic species originate from windblown mineral dust and ocean salts from wave breaking and bubble bursting (Poschl and Shiraiwa 2015). Sulfur dioxide gas emitted in volcanic eruptions is transformed into secondary sulfate salts, while many other sources co-emit volatile organic carbon compounds (VOCs) as gasses (Poschl and Shiraiwa 2015). Atmospheric oxidants continuously transform VOCs into less volatile secondary species that can then partition

from the gas phase to the condensed phase (Poschl and Shiraiwa 2015). It has been observed that inorganics like sulfuric acid are important in new particle formation in combination with smaller organic acids (Zhang 2010, Riccobono et al. 2014, Bianchi et al. 2016), and that the bulk of these conversions from the gas phase form as organic coatings on the surface of pre-existing aerosol (Riva et al. 2017, Zelenyuk et al. 2017). Organic species formed and transformed by atmospheric processes are known as *secondary organic aerosol*. The observed mass fraction of secondary organic aerosol is dominant compared to primary emissions, due to the oxidizing nature of the atmosphere (Paglione et al. 2014, Poschl and Shiraiwa 2015, Gilardoni et al. 2016, Shrivastava et al. 2017). Unstable molecules like ozone (O_3) and hydroxyl radicals are highly reactive with reduced carbon species, especially unsaturated double bonds and ring structures. Photolysis reactions in direct sunlight can also occur for molecules which absorb solar radiation and photolysis reactions are the driving force for the formation of ozone and hydroxyl radicals (Kroll and Seinfeld 2008). Nitrogen oxides (NO and NO_2), commonly known as NO_x , are anthropogenic pollutants that play a part in many atmospheric reaction pathways and thus affect the oxidation of aerosol (Kroll and Seinfeld 2008, Shrivastava et al. 2017). At night, nitrate radicals ($NO_3^{\bullet}$) also play a role in secondary organic aerosol formation (Boyd et al. 2015, Shrivastava et al. 2017). Aerosol particles and water-soluble gasses can be dissolved in the aqueous phase of cloud and fog droplets, where they are further transformed in an altogether different environment than the gas phase (Darer et al. 2011, Zhao et al. 2013, Boone et al. 2015, El-Sayed et al. 2015). Furthermore, aerosol can begin to absorb water at relative humidity levels below supersaturation conditions and create a layer of *aerosol liquid water*, which is another distinct reaction environment that can have high concentrations of acids, salts and organics (McNeill 2015, Sareen et al. 2017). As these different forms of liquid water evaporate newly transformed aqueous phase processed aerosol are left behind (Lee et al. 2013, Ervens 2015). Because of these wide variety of sources and transformation processes atmospheric aerosol exists as a mixture with extreme chemical complexity, particularly for organics.

Primary biological aerosol include pollen, fungal spores, algae, lichens, insect brochosomes and eggs, plant and animal fragments (hair, dander, debris), viruses, bacteria, and bacterial fragments (Despres et al. 2012). Trees and other plants secrete volatile oils, such as isoprene, monoterpenes and sesquiterpenes, which are the largest emissions of VOCs globally and thus the largest source of SOA on Earth (Goldstein and Galbally 2007). All forms of combustion suffer from some inefficiency in the progression from hydrocarbon fuels to CO_2 (Wielgosinski 2012), and thus anthropogenic sources of aerosol include motor-vehicle and industrial exhaust (Rinaldi et al. 2007, Bahreini et al. 2009, Tong et al. 2016). Biomass combustion is another significant source of aerosol and VOCs to the atmosphere (Bond et al. 2004), including wildfires, agricultural land clearing and residential combustion for heating and cooking (Andreae and Merlet 2001, Reid et al. 2005, Mazzoleni et al. 2007, Fleming et al. 2018). Fuels can include wood, dried vegetative materials and animal dung (Reid et al. 2005, Fleming et al. 2018).

Figure 1.1. A hypothetical chemical structure for the biopolymer lignin (a) which is composed mainly of monomer units of sinapyl alcohol (b), coniferyl alcohol (c) and p-coumaryl alcohol (d). During the combustion of lignin, the main chemical byproducts are the derivatives of these three monomer units, syringol (e), guaiacol (f) and phenol (g) as well as further substituted forms.

There are many tracer species associated with biomass burning emissions especially for wood combustion. Anhydrosugars like levoglucosan and its stereoisomers are produced from the decomposition of cellulose and hemicellulose (Simoneit 2002, Mazzoleni et al. 2007). Methoxyphenols and other substituted phenols are produced from the decomposition of the biopolymer lignin (see Fig. 1.1), which makes up approximately 30% of woody tissue by mass (Simoneit 2002, Mazzoleni et al. 2007). The types of combustion byproducts in biomass burning are highly dependent on the type of fuel, moisture content, the size of the fire, oxygen availability and the burning phase of the fire (Reid et al. 2005). Flaming fires have higher temperatures and higher combustion efficiency, where PAHs and small particles of black carbon can be produced as carbon radical species are formed from broken bonds (Lima et al. 2005, Reid et al. 2005). Smoldering fires produce more smoke and larger particulates due to lower temperatures and lower combustion efficiency (Reid et al. 2005). The intense heat from large fires can also facilitate long range transport of biomass burning emissions (Damoah et al. 2004, Yu et al. 2019), as they are lofted into higher parts of the atmosphere where colder and dryer conditions may inhibit atmospheric aging due to viscosity changes (Hinks et al. 2016, Reid et al. 2018, Schum et al. 2018).

Until recently organic aerosol was thought to be present exclusively in the liquid state, but it is now known that semi-solid and solid states are possible (Reid et al. 2018). The viscosity of aerosol particles plays a role in the oxidation rate by gas phase components like ozone, which are less likely to permeate into solid-phase particles (Hinks et al. 2016). A solid aerosol phase state may help to explain the observation of semi-volatile and reduced carbon organic pollutants in remote areas following long range transport (Shrivastava et al. 2017, Schum et al. 2018). Semi-solid and solid phases are more likely in the higher altitudes of the free troposphere, where conditions are colder and dryer (Shiraiwa et al. 2017, Schum et al. 2018). Hygroscopic aerosol particles begin to absorb water vapor at relatively low humidity and condensed liquid water can plasticize aerosol into semi-solid and liquid states (Reid et al. 2018). The individual chemical components can react more easily in the aqueous phase where concentrations can be notably higher depending on liquid water content, enhancing the reaction rates and types of transformations which can occur (Ervens et al. 2011, Sareen et al. 2017). In fog and cloud droplets, hydrolysis reactions are enhanced due the relatively high amount of water molecules (Darer et al. 2011, Hu et al. 2011). Aqueous phase reactions can be influenced by pH, which tends to be low in aerosol due to dissolved nitric, sulfuric and organic acids (Giulianelli et al. 2014, Herckes et al. 2015). However, ammonia emissions from farming and animal husbandry can increase pH, especially in the absence of acid precursors like sulfur dioxide (SO₂) (Warner et al. 2017). Dissolved metal ions can catalyze oxidation through OH[•] production and lead to chemical transformation and further complexity (Deguillaume et al. 2005, Holmes and Petrucci 2007). Aerosol processed by cloud droplets tends to be more oxidized than those that have been transformed in the gas phase and the presence of oligomeric compounds formed in SOA are also presumed to be from aqueous phase reactions (Ervens et al. 2011). As all fog and cloud droplets are seeded by some pre-existing aerosol, this makes each droplet a potential reaction vessel for the formation of secondary organic aerosol, as well as potential sink for aerosol through wet deposition and precipitation.

1.1.4 Aerosol sampling

Many aerosol studies use online ambient monitoring techniques which record data in real time. These instruments are often located in remote areas (Fialho et al. 2005), urban centers (Chow et al. 2002) or even mounted in aircrafts to study moving plumes and high-altitude variations (Shinozuka et al. 2007, Sedlacek et al. 2018). These methods excel at measuring the optical properties, size distributions and number concentrations of aerosol. Real time chemical characterization of aerosol is challenging, though field deployable aerosol mass spectrometers are capable of chemical characterization with high time resolution (Drewnick et al. 2005, DeCarlo et al. 2006, Aiken et al. 2008, Gilardoni et al. 2016). However, these analyses are somewhat limited due to the high fragmentation of analyte molecules (see section 1.2.2) and require complex deconvolution of spectra to interpret results (DeCarlo et al. 2006). For more detailed chemical characterization, atmospheric samples must be collected for subsequent offline laboratory analysis.

Sampling atmospheric aerosol can be done in a variety of ways. Most commonly samplers pull streams of air at approximately 30 to 100 m³ h⁻¹ and collect aerosol particles by impaction onto quartz fiber or polytetrafluoroethylene (PTFE) filters (Mazzoleni et al. 2012, Kourchev et al. 2014, Dzepina et al. 2015). PM_{2.5} mass concentrations typically range from 10 to 100 µg m⁻³ (Lee et al. 2005, Weimer et al. 2006, Rinaldi et al. 2007, Costabile et al. 2017) with more concentrated episodes being considered severe pollution. For this reason, large volumes of air must be sampled for 12 to 24 hours to obtain enough aerosol material for experimental analysis. Ambient aerosol experiments are often sample limited, meaning many studies have focused on just one or two analytical methods. Liquid phase samples of cloud and fog water can be obtained with samplers that pull ambient air through a chamber lined with multiple series of inert strands, where the droplets collect and coalesce on the strands and then flow into a collection bottle (Fuzzi et al. 1997). Collected filters and liquid phase samples are usually stored frozen until analysis can be performed to minimize the degradation of chemical species post collection.

1.2 Introduction to mass spectrometry

1.2.1 Molecular mass

Mass spectrometry is an analytical method for the detection of ionized molecules. These extremely sensitive instruments are capable of trace level detection and analysis of samples in low concentration. The measured mass-to-charge ratio (m/z) of detected ions allows for the determination of the mass of an individual molecule when the charge of the ion is known (Gross 2004, Hoffman and Stroobant 2007). The conventional unit for mass used in mass spectrometry is Daltons (Da) (Gross 2004). Some clarity is required when discussing the masses obtained in mass spectrometry, as the masses of the atoms which make up a molecule can be defined in several ways with subtle differences. The *average mass* of an atom which is reported on the periodic table, accounts for all of the naturally occurring isotopes of an element and the respective abundances (Gross 2004). As the

different isotopes of atoms differ by one or more Da they are easily resolved in mass spectrometry measurements, thus the average mass is not often discussed. The *nominal mass* is the integer mass of an atom assuming each proton and neutron of the atom have a mass of exactly 1 (Gross 2004). The *monoisotopic mass* is the exact mass of the most abundant naturally occurring isotope of atom (Gross 2004, Hoffman and Stroobant 2007). Monoisotopic masses are defined relative to carbon, with the monoisotopic mass of ^{12}C equaling exactly 12 (Gross 2004). The calculated difference between monoisotopic mass and nominal mass describes *mass defect*, which can be used to describe molecular trends (Gross 2004). These distinct definitions of mass are discussed in Table 1.1 for a set of elements specific to this work. Atoms with nominal mass greater than their monoisotopic mass like ^{16}O are described as mass deficient, while atoms with nominal mass less than their monoisotopic mass like ^1H are described as mass sufficient (Gross 2004). The monoisotopic masses defined for ^{12}C are set by the International Union of Pure and Applied Chemistry (IUPAC). However, monoisotopic masses can be renormalized using some other nominal and monoisotopic mass pair, which results in the *Kendrick mass* (Hughey et al. 2001, Gross 2004). Kendrick mass is discussed in detail in later chapters, but most commonly in this work a Kendrick mass based on a unit of CH_2 is used, renormalizing measured monoisotopic masses to 14. The resulting difference between Kendrick mass and nominal mass describes a specific type of mass defect known as the *Kendrick mass defect* (Hughey et al. 2001, Gross 2004).

Table 1.1 A brief summary of elemental isotopes relevant to atmospheric aerosol and their different mass values.

Isotope	Nominal mass	Monoisotopic mass	Mass defect
^{12}C	12	12.00000	0.00000
^{13}C	13	13.00335	+0.00335
^1H	1	1.007825	+0.007825
^{14}N	14	14.00307	+0.00307
^{16}O	16	15.99491	-0.00509
^{32}S	32	31.97207	-0.02793
^{34}S	34	33.96787	-0.03213

1.2.2 Ionization of analyte molecules

Analyte molecules are detected by incident electrical currents in mass spectrometer instruments, thus the analytes must be charged ions to be detected (Hughey et al. 2001, Gross 2004, Pavia 2009). Ions can also be easily separated from neutral species using electrostatic fields, which can further be used to facilitate movement through the internal paths of the instruments (Gross 2004, Hoffman and Stroobant 2007). Masses can be measured in both the positive and negative polarities, allowing for versatility in the detection of desired analytes. Impact by high energy electrons (70 eV) is a traditional ionization method in mass spectrometry (Gross 2004, Hoffman and Stroobant 2007, Pavia 2009). This method leads to excessive yet reproducible fragmentation of the analyte

molecules, making chemical characterization possible through the comparison to spectral libraries (Gross 2004). The fragmentation caused by *hard-ionization* methods like electron impact is not ideal for the analysis of complex mixtures, as the resulting spectra are convoluted and difficult to interpret. *Soft-ionization* techniques result in minimal fragmentation and the detection of molecules which are essentially intact, save the abstracted or adducted protons which promote charge (Gross 2004, Hoffman and Stroobant 2007). The detection of intact molecules is especially important for the assignment of molecular formulas, as described in this work. These soft-ionization techniques typically produce ions with a charge of ± 1 . Though multiply charged ions are possible, especially for larger molecules (Konermann et al. 2013) based on the observed m/z differences in ^{12}C and ^{13}C species in this work we do not expect significant influence from ions with a charge greater than 1.

Common soft-ionization techniques include atmospheric pressure chemical ionization (APCI), atmospheric pressure photo-ionization (APPI) and electrospray ionization (ESI) (Gross 2004, Hoffman and Stroobant 2007). In APCI, a corona discharge is used to generate primary charged ions of a reagent gas (commonly N_2), which then transfers charge through molecular collisions with the nebulized analyte molecules (Hoffman and Stroobant 2007). In APPI, photons from a discharge lamp are used to ionize the nebulized analytes, often with the use of a dopant molecule (such as toluene) which will absorb the UV light emitted by the lamp and transfer charge to the analytes (Kauppila et al. 2002). In ESI, solvent containing analytes flows through a charged capillary tube, where charge gathers at the solvent surface (Konermann et al. 2013). At the end of the capillary the solvent is dispersed by a coaxial gas (commonly N_2), creating a spray of small highly charged droplets, which repel one another as the solvent dries, creating individual ions (Konermann et al. 2013). These three ionization techniques have their own strengths and weaknesses regarding analyte detection, which are discussed briefly here and in greater detail in later chapters. ESI is best for the detection of large molecules ($> 10,000$ Da) like proteins and polymer compounds as well as the detection of small ($< 1,000$ Da) polar and semi-polar molecules (Hoffman and Stroobant 2007, Konermann et al. 2013). APCI and APPI allow for the ionization less polar molecules than with ESI, and APPI even more so for molecules which are highly aromatic and thus more likely light absorbing (Kauppila et al. 2002, Hoffman and Stroobant 2007).

Regarding complex mixtures of organics, ESI in the negative ion mode is a common analytical technique (Schmitt-Kopplin et al. 2010, LeClair et al. 2012, Dzepina et al. 2015, Cook et al. 2017) and allows for the detection of several thousand different monoisotopic masses depending on sample (see Figs. 1.2 and 1.3). However, in ESI molecules compete for charge during ionization, meaning the detected abundance of an ion is influenced by both the analyte concentration as well as its ionization efficiency (Cech and Enke 2001, Hoffman and Stroobant 2007, Konermann et al. 2013). For this reason, ESI is a non-quantitative technique. This also means that there are potential biases introduced to the analysis of complex mixtures while using ESI. As molecules compete for charge in the ESI source, molecules with a higher ionization efficiency may gather charge more quickly and effectively suppress the detection of other molecules in the mixture. Avoiding these charge

competition biases can be challenging but there have been successful studies of atmospheric aerosol samples that use alternate ionization and chromatography techniques (Winterhalter et al. 2003, Desyaterik et al. 2013, Lin et al. 2018). The separation of analytes by chromatography could isolate those species which are highly efficiently ionized from those which are not and allow for the detection of less efficiently ionized molecules.

1.2.3 Ultrahigh resolution mass spectrometry

Complex organic mixtures like atmospheric aerosol contain a variety of functional groups and heteroatoms which make their molecular characterization challenging. Mathematically, a molecular formula ($C_cH_hN_nO_oS_s$) can easily be assigned to a measured mass, by substituting the corresponding exact monoisotopic masses for the elements contained in the formula. However, the quality of data acquired by mass spectrometers needs to be considered regarding the assignment of molecular formulas. For example, when considering only the nominal mass of organic molecules, it is impossible to distinguish between mass moieties such as O_2 vs. S or C_2H_4 vs. N_2 as they have the same nominal mass. With nominal mass measurements molecular formulas cannot be assigned to mass spectrometry data with any degree of certainty because of this ambiguity. However, as resolving power between the detected peaks increases, more and more confidence can be attributed to the measured monoisotopic masses of the analytes. Quadrupole and linear ion trap instruments are capable of nominal mass resolving powers of $\sim 2,000$ to $4,000$, but with extremely fast scan times (Hoffman and Stroobant 2007). Some high resolution time-of-flight instruments are capable of resolving powers of $\sim 20,000$ (Hoffman and Stroobant 2007) and when combined with prior chromatographic separation are capable of reliable molecular formula assignments (Desyaterik et al. 2013). Ultrahigh resolution instruments like the two used in this work, the Fourier transform-ion cyclotron resonance (FT-ICR) and Orbitrap mass spectrometers, are capable of resolving powers of $200,000$ to $400,000$ and beyond (Marshall 2000, Zubarev and Makarov 2013). These instruments avoid the ambiguity in mass moieties with the same nominal mass discussed earlier, by exploiting their very small differences in monoisotopic mass. For example, O_2 vs. S has a monoisotopic mass difference of 17.76 mDa, and C_2H_4 vs. N_2 has a difference of 25.15 mDa. Resolving masses with these minute differences is extremely important for the reliable assignment of molecular formulas, especially when considering the inclusion of heteroatoms like N and S. An example of the enhanced peak resolution provided by ultrahigh resolution instruments is provided in Fig. 1.2. The aerosol sample analyzed by negative mode ESI in Fig. 1.2 shows only one broad peak at m/z 436 when analyzed with a linear ion trap, but this resolves into several peaks at $60,000$ resolving power with the Orbitrap. Further increases in resolving power above $60,000$ show additional peaks with resolutions of $120,000$ and $240,000$, with clear baseline separation with resolutions of $240,000$ and $480,000$. The expected reduced ion signal at higher resolving power from longer scan times can also be observed in Fig. 1.2 with the decreasing relative abundance of peaks as resolving power increases.

Figure 1.2. The profile mass spectra collected for a biomass burning influenced the biomass burning influenced ambient aerosol sample BB06 (see chapter 3) which was analyzed using negative mode electrospray ionization. Linear ion trap (IT) and Orbitrap (60,000 to 480,000) mass analyzers were used and excerpts of the spectra collected are arranged from top to bottom panels by increasing resolving power. The metric prefix kilo- (1,000) is used the abbreviate these resolving powers.

Ultrahigh resolution instruments allow for the straightforward direct infusion analysis of complex mixtures with minimal sample preparation and analysis time. However, these types of analysis are not capable of distinguishing between molecular isomers. Thus, no chemical structural identities can be assumed from the detected masses without some additional analytical method such as tandem mass spectrometry (MSⁿ) or chromatography. Long scan times makes coupling these instruments to chromatography challenging, though chromatographic fractions can be collected ahead of time and then analyzed individually by direct infusion.

1.2.3.1 Fourier transform-ion cyclotron resonance instruments

FT-ICR instruments use a magnetic field (~7 to 21 T) to trap analyte ions in a cyclotron orbit within the mass analyzer (Hoffman and Stroobant 2007). These orbits have a specific frequency, related to the analyte mass. Ions are detected through excitation by a resonant electromagnetic wave which increases the radius of the orbit inducing a measurable difference in voltage between the walls of the mass analyzer (Hoffman and Stroobant 2007). A Fourier transform calculation is then used to interpret the mass spectra from these measured changes in electrical potential (Hoffman and Stroobant 2007). FT-ICR instruments are capable of resolving powers above 1,000,000 though these elevated resolutions require longer scan times which may lead to the decay of ion signal. While FT-ICR instruments have some of the highest available resolving powers, these instruments are quite costly to support as they require a steady supply of cryogenic He and N₂ gasses to maintain the superconducting magnet.

1.2.3.2 Orbitrap instruments

Orbitrap instruments use an electrostatic field to trap analyte ions around a spindle shaped central electrode (~3 kV) (Hoffman and Stroobant 2007). Ions are injected into the trap at an angle with high kinetic energy and thus orbit the central electrode while also oscillating along the spindle creating intricate spiral patterns (Hoffman and Stroobant 2007, Zubarev and Makarov 2013). These oscillations induce a current which can be measured and converted to a mass spectrum following a Fourier transform calculation (Hoffman and Stroobant 2007). While early generation Orbitrap instruments were capable of resolving powers of 60,000 to 100,000 modern instruments are capable of resolving powers of 1,000,000 (Denisov et al. 2012). Specifically, the Orbitrap Elite instrument located at the Chemical Advanced Resolution Methods (ChARM) Laboratory at Michigan Technological University is capable of resolving powers of 120,000 to 480,000. As with FT-ICR instruments, these higher resolving powers require longer scan times and can decrease ion signal, especially at higher m/z . Orbitrap instruments present an economical alternative to FT-ICR instruments with equal capabilities in resolving power.

1.2.4 Sample preparation

Aerosol filters can be extracted in water (Schmitt-Kopplin et al. 2010, LeClair et al. 2012, Dzepina et al. 2015, Cook et al. 2017) or with the use of organic solvent mixtures (Putman et al. 2012, Lin et al. 2018) and the solvent used plays a strong role in what chemical components are observed during analysis (Tfaily et al. 2015). Water extracts can be combined with a solid phase extraction (SPE) method used to exchange to an organic solvent mixture and concentrate the analytes, as solvent compositions high in water evaporate less efficiently in the ion source and can cause poor ion signal (Cech and Enke 2001). Analytes dissolved in water are loaded onto a conditioned SPE cartridge and captured by the stationary phase of the cartridge. The aqueous phase then flows through the cartridge and the analytes can be eluted with an organic solvent. Some fraction of analytes is inexorably lost in this process, as some polar species are eluted with the aqueous phase having never been retained by the stationary phase, and some non-polar species are never eluted off the stationary phase by the organic solvent. Additional tandem cartridges with alternate chemistry can be employed to capture those polar compounds which are unretained by the first column, and additional elution with stronger organic solvents can be used to separate those non-polar species which remain on the cartridge. However, these additional steps each add a layer of difficulty to sample preparation and the interpretation of results and are uncommon in the literature for atmospheric aerosol. Other more sophisticated analytical techniques, like nanospray desorption ESI, perform solvent extraction of the aerosol filter and direct infusion into the mass spectrometer simultaneously (Laskin et al. 2012, Fleming et al. 2018).

Simple organic solvent extraction of aerosol filters, such as with acetonitrile, can avoid the complications of SPE. However, the volume of solvent used in extraction greatly effects the concentration of the extracted analytes and large aerosol filters may require large volumes of solvent to fully extract the filter. The extracts can be evaporated with streams on N₂ gas to concentrate the analytes, however this may lead to the loss of some more volatile analyte species. These simple organic extractions are also not viable for liquid phase collected samples like fog and cloud water, and SPE must be used in sample preparation.

1.2.5 Post processing and data analysis

As stated previously, the organic component of atmospheric aerosol is an extremely complex mixture and collected mass spectra can contain thousands of peaks above the signal-to-noise threshold (Schmitt-Kopplin et al. 2010, Mazzoleni et al. 2012, Dzepina et al. 2015, Cook et al. 2017). As an example of this complexity, the reconstructed mass spectra for the assigned molecular formulas of a biomass burning aerosol sample is shown in Fig. 1.3. Through the progressively smaller scan ranges shown in the Figure, it can be observed that each nominal mass contains several different molecular species. Though the presence of isomers cannot be definitively proven with this data it has been shown for similar mixtures of natural organic matter that each molecular formula likely represents

multiple different isomeric species, on a scale of 1 to 2 orders of magnitude more isomers than the number of observed molecular formulas (Zark et al. 2017, Hawkes et al. 2018).

The assignment of molecular formulas is a two-step process, the assignment itself and the meticulous validation of the quality of the formulas through chemical trends. This process is outlined in general terms in the following two subsections, as well as in each additional chapter with the details specific to those respective data sets.

1.2.5.1 Molecular formula assignment

Molecular formula assignment software is used to calculate possible molecular formulas for a given mass. In this work two different software tools were used, “Composer” (Sierra Analytics v. 1.0.5) and an in house developed R package “MFAssignR” (Schum et al. 2018). Both tools operate on a similar set of chemical principles. Though the internal mechanics of each tool may or may not rely on these specific principles, the science behind them are upheld in the final result. The molecular formula calculators used are able to exploit extremely small, yet measurable, differences in monoisotopic masses to determine acceptable molecular formulas. Because of the extremely small margins of error in these types of calculations, exceedingly accurate mass measurements are required for reliable formula assignment.

It is important to employ additional measures to ensure proper assignments for masses at higher molecular weights as the number of mathematically possible formulas increases drastically at higher masses. Kendrick mass and Kendrick mass defect for a CH₂ homologous series is often used to relate molecular formulas of natural organic matter systems and extend assignments beyond a certain mass threshold. Molecular formulas with the same Kendrick mass defect fall into *homologous series* with the same number of heteroatoms and double bond equivalents (DBE), meaning the formulas only differ by units of CH₂ (Stenson et al. 2003). The limit in mass allowed during formula assignment is known as a *de novo* cut off and typically an *m/z* of 350 to 500 is used. This process gives additional credibility to the molecular formulas calculated for ions of higher molecular weights, as no formulas will be assigned above this limit without belonging to a homologous series that begins below the limit. Please note that these relationships between molecular formulas of a CH₂ homologous series do not correspond to relationships in chemical structure.

Typically, C, H and O atoms are allowed with relatively no restrictions, while heteroatoms like N, S and P are severely limited to three or less per molecule (Mazzoleni et al. 2012, Dzepina et al. 2015, Cook et al. 2017, Schum et al. 2018). These restrictions help to reduce the overall number of unfeasible and unexpected molecular formulas which are only mathematically possible. For the same reason the isotopic analogs of these species are not allowed as they are only expected in low abundance. These limitations help to prevent poor formula assignments from entering the data set and overall aid in the validation process.

Figure 1.3. Reconstructed mass spectra for the biomass burning influenced ambient aerosol sample BB06 (see chapter 3) which was analyzed by Orbitrap Elite mass spectrometry using negative mode electrospray ionization. Spectra are inset to progressively smaller scan ranges to demonstrate the typical level of complexity for the mass spectra of atmospheric aerosol samples. The full spectrum from m/z 100 to 800 contains 5,415 assigned molecular formulas, while the first inset from m/z 200 to 300 contains 1,069 formulas, the second inset from m/z 275 to 285 contains 122 formulas and the final inset from m/z 280.9 to 281.3 contains 20 formulas.

1.2.5.2 Molecular formula validation

An extensive quality assurance process is used to validate the assigned molecular formulas and remove those which are most likely to be incorrectly assigned. The *rule of 13* describes a base hydrocarbon molecular formula (C_nH_{n+r}) for any measured mass by solving equation (1.1):

$$\left(\frac{mass}{13}\right) = n + \left(\frac{r}{13}\right) \quad (1.1)$$

Where n is an integer quotient, and r is the remainder (Pavia 2009). Along similar lines, this base formula also defines the number of heteroatoms (O, N and S) allowed in the final formula, as each heteroatom will require some substitution of C_xH_y to maintain the same measured mass. This base hydrocarbon formula can be used to determine the maximum number of heteroatoms, as well as the minimum number of hydrogen atoms, in an acceptable molecular formula.

The absolute error (AE) of a calculated molecular formula can be calculated from the measured mass of the ion and the theoretical mass of the assigned molecular formula using equation (1.2):

$$AE \text{ (ppm)} = \left| \frac{measured\ mass - theoretical\ mass}{theoretical\ mass} \right| * 10^6 \quad (1.2)$$

AE is typically expressed in parts per million (ppm), and commonly AE values are restricted to values of 2 to 3 ppm. Though isotopes are typically not allowed during formula assignment, the presence or absence of isotopic species at the corresponding mass to an assigned formula can be used as a method of validation. Isotope identification is especially useful for ^{13}C and ^{34}S as these are naturally abundant enough to be detected. Other mathematical trends can be used to validate the molecular formula assignments based on the corresponding number of atoms in the assigned formula. The DBE of a formula can be used to estimate the number of double bonds and ring structures present in formula and is calculated using equation (1.3) for the format $C_cH_hO_oN_nS_s$:

$$DBE = c - \left(\frac{h}{2}\right) + \left(\frac{n}{2}\right) + 1 \quad (1.3)$$

Molecular formulas are required to have integer values for the calculated DBE, with an upper limit of approximately 30 to 40. Similarly, hydrogen saturation and oxygenation follow predictable trends and ratios of the number of these atoms with respect to carbon can help to maintain reliable formulas. O:C ratios are limited to 2 and H:C ratios are limited to 3 as organic structures with ratios higher than this are not expected. The *nitrogen rule* uses the odd numbered valence of nitrogen to determine an acceptable number of nitrogen in a formula, as molecular formulas with odd masses must also contain an odd number of nitrogen atoms and even masses must therefore contain an even number (or 0) of nitrogen atoms (Pavia 2009). Trends in the difference between DBE and oxygen number (DBE-O)

are described in Herzsprung et al. (2014) where reliable formula assignments have a DBE-O value between -13 and 13. Molecular formulas which are part of a homologous series created by Kendrick mass defect which have large gaps between series or large gaps within series likely include formulas which are incorrectly assigned.

The preceding validation methods are an uncomprehensive overview of a crucial process in this work and help to ensure that the highest quality of data is obtained for interpretation.

2 Molecular insights on aging and aqueous phase processing from ambient biomass burning emissions-influenced Po Valley fog and aerosol¹

2.1 Abstract

To study the influence of regional biomass burning emissions and secondary processes, ambient samples of fog and aerosol were collected in the Po Valley (Italy) during the 2013 Supersito field campaign. After the extent of “fresh” vs. “aged” biomass burning influence was estimated from proton nuclear magnetic resonance (¹H-NMR) and high resolution time of flight aerosol mass spectrometry (HR-ToF-AMS), two samples of fog water and two samples of PM₁ aerosol were selected for ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) analysis. Molecular compositions indicated that the water-soluble organic matter was largely non-polymeric without clearly repeating units. The selected samples had an atypically large frequency of molecular formulas containing nitrogen and sulfur (not evident in the NMR composition) attributed to multifunctional organonitrates and organosulfates. Higher numbers of organonitrates were observed in aerosol, and higher numbers of organosulfates were observed in fog water. Consistent with the observation of an enhanced aromatic proton signature in the ¹H-NMR analysis, the average molecular formula double bond equivalents and carbon numbers were higher in the “fresh” biomass burning influenced samples. The average O:C and H:C values from FT-ICR MS were higher in the samples with an “aged” influence (O:C = 0.50 to 0.58 and H:C = 1.31 to 1.37) compared to those with “fresh” influence (O:C = 0.43 to 0.48 and H:C = 1.13 to 1.30). The “aged” fog had a large set of unique highly oxygenated CHO fragments in the HR-ToF-AMS, which reflects an enrichment of carboxylic acids and other compounds carrying acyl groups, highlighted by the NMR analysis. Fog compositions were more oxidized and “SOA-like” than aerosols as indicated by their NMR measured acyl vs. alkoxy ratios and the observed molecular formula similarity between the aged aerosol and fresh fog, implying that fog nuclei must be somewhat aged. Overall, functionalization with nitrate and sulfate moieties, in addition to aqueous oxidation, trigger an increase in the molecular complexity in this environment, which is apparent in the FT-ICR MS results. This study demonstrates the significance of the aqueous phase to transform the molecular chemistry of atmospheric organic matter and contribute to secondary organic aerosol.

¹As stated in the Preface, the material contained in chapter 2 was previously published in the journal *Atmospheric Chemistry and Physics*:

Brege, M., Paglione, M., Gilardoni, S., Decesari, S., Facchini, M.C. and Mazzoleni, L.R. (2018). "Molecular insights on aging and aqueous-phase processing from ambient biomass burning emissions-influenced po valley fog and aerosol." *Atmospheric Chemistry and Physics* 18(17): 13197-13214.

This article is © 2018 by the authors, and licensed for distribution by Copernicus Publications, Göttingen, Germany. The works published in this journal are distributed under the Creative Commons Attribution 4.0 License. See Appendix B for documentation of permission to republish the content from this article.

2.2 Introduction

Atmospheric organic aerosol particles are comprised of a complex mixture of numerous individual organic compounds, produced by direct emissions and secondary processes, of which a significant impact is from transformations in the aqueous phase. Surface emitted primary organic aerosol and volatile organic compounds are transformed in the atmosphere by gas to particle phase conversion, heterogeneous reactions, and aqueous phase reactions in aerosol water, fog, and cloud droplets (Ervens et al. 2011, Herrmann et al. 2015). The products of these processes are collectively referred to as secondary organic aerosol (SOA). These aging reactions happen quickly in the atmosphere, and the observed mass fraction of SOA is larger than that of primary organic aerosol (Zhang et al. 2007, Ervens et al. 2011, Zhang et al. 2011, Paglione et al. 2014, Gilardoni et al. 2016). Biomass burning emissions, such as those from forest fires, agricultural land clearing, residential heating, and cooking with biofuels, are important sources of organic carbon to the atmosphere globally (Andreae and Merlet 2001, Bond et al. 2004, Glasius et al. 2006, Laskin et al. 2015). Biomass burning products include simple organic acids, sugars and anhydrosugars, substituted phenols, polycyclic aromatic hydrocarbons, and other compounds, depending on the type of fuel and burn conditions (Mazzoleni et al. 2007, Pietrogrande et al. 2014, Pietrogrande et al. 2014, Gilardoni et al. 2016). These water-soluble emissions can serve as precursors for SOA once dissolved in the aqueous phase (Chang and Thompson 2010, Yu et al. 2014, Yu et al. 2016), and upwards of 50% of organic matter in fog and cloud droplets remains unidentified (Herckes et al. 2013). Biomass burning emissions can even facilitate droplet nucleation. In fact, laboratory studies indicate that in addition to hydrophilic species, even refractory “tar balls,” emitted from smoldering biomass burning begin to absorb water at high relative humidity (Hand et al. 2005, Laskin et al. 2015).

Atmospheric chemistry models are currently unable to replicate several key aspects of SOA, including SOA concentration levels, chemical oxidation states, degree of functionalization, and the occurrence of high molecular weight compounds, such as atmospheric humic-like substances (Ervens et al. 2011, Lee et al. 2013, Nguyen et al. 2013). Aqueous phase reactions in wet aerosol, cloud, and fog droplets have been proposed to improve these SOA observation gaps (Ervens et al. 2011, Herckes et al. 2013, Laskin et al. 2015, Gilardoni et al. 2016), but the current level of understanding regarding aqueous phase processes is insufficient to include them in models. Laboratory studies focusing on simplified systems of only one or two precursor components have successfully recreated some of the complexity of ambient atmospheric samples (De Haan et al. 2011, Lee et al. 2013, Nguyen et al. 2013, Hawkins et al. 2016, Yu et al. 2016). A number of recent studies focusing on the molecular composition of cloud (Lee et al. 2012, Desyaterik et al. 2013, Pratt et al. 2013, Zhao et al. 2013, Boone et al. 2015, Cook et al. 2017) and fog (Mazzoleni et al. 2010, LeClair et al. 2012, Xu et al. 2017) chemistry have been recently reported. Together these studies indicate a clear importance of aqueous phase reactions for the production of aqueous SOA, including the formation of organonitrates, organosulfates, and nitrooxy-organosulfates. Of these, organosulfate formation is thought to happen nearly exclusively in the aqueous phase (Ervens et al. 2011, Herrmann et al. 2015). Along with organonitrates, organosulfates are susceptible to hydrolysis in the aqueous phase, though

high kinetic barriers under atmospheric conditions often slow these reactions and allow for the observation of these species in ambient samples (Darer et al. 2011, Hu et al. 2011). Organosulfates are often described in the literature as the products of acid catalyzed oxidation of biogenic terpenoids (Surratt et al. 2008, Pratt et al. 2013, Schindelka et al. 2013), but have also been observed in biomass combustion influenced cloud water (Zhao et al. 2013, Cook et al. 2017). The formation of aqueous phase products in aerosol, fog and cloud waters, greatly increase the complexity of organic aerosol. Although several analytical techniques have been used to address the challenge of resolving the complex mixture of atmospheric organic matter (Decesari et al. 2007, Hertkorn et al. 2007, Nizkorodov et al. 2011, Desyaterik et al. 2013, Dall'Osto et al. 2015, Noziere et al. 2015, Laskin et al. 2016, Willoughby et al. 2016), no universal analytical method exists.

The Po Valley (Italy) has ideal ambient conditions to study aqueous phase influences on atmospheric organic matter. The valley contains a mixture of densely populated areas and intensively cultivated agricultural regions. Surrounded by mountains to the north, west and south, the valley frequently has stable meteorological conditions with low ventilation and a low boundary layer, allowing for the accumulation of high concentrations of regional pollutants. Consequently, frequent fog events and high concentrations of anthropogenic biomass burning emissions are observed in months with cold temperatures (Larsen et al. 2012, Saarikoski et al. 2012, Giulianelli et al. 2014, Paglione et al. 2014, Gilardoni et al. 2016). The Po Valley has some of the highest reported carbon concentrations for fog water in the world (Herckes et al. 2013). In recent years, the analysis of fog water and aerosol from San Pietro Capofiume (SPC) (located 30 km northeast of the city of Bologna) has included Aerodyne high resolution time-of-flight aerosol mass spectrometry (HR-ToF-AMS) and proton nuclear magnetic resonance spectroscopy (^1H -NMR) to determine the fog scavenging efficiency of aerosol (Gilardoni et al. 2014) and source apportionment of aerosol (Decesari et al. 2007). In Saarikoski et al. (2012), HR-ToF-AMS data from SPC aerosol showed an extremely high concentration of aerosol nitrate (39%) and a somewhat typical concentration of organic carbon (33%) in agreement with Gilardoni et al. (2014). Positive matrix factorization (PMF) of HR-ToF-AMS organic mass fragments was used to identify several factors describing Po Valley organic aerosol, including factors for fresh biomass burning organic aerosol, and three types of oxygenated organic aerosol (Saarikoski et al. 2012). A similar study by Paglione et al. (2014) used PMF on ^1H -NMR data of SPC aerosol to identify factors for fresh biomass burning emissions, as well as SOA factors, including products formed from aged biomass burning emissions.

Further investigation with a focus on molecular markers and source apportionment was done as part of the Supersito 2013 field campaign in the Emilia-Romagna region, including samples from SPC and the urban site of Bologna (Pietrogrande et al. 2014, Pietrogrande et al. 2014, Poluzzi et al. 2015). The campaign has shown the significance of biomass burning emissions in the region. Approximately 35% of the organic carbon was from wood burning in winter months (Pietrogrande et al. 2015), biomass burning emissions were shown to increase with decreasing ambient temperature (Gilardoni et al. 2014), and aqueous phase SOA formation from biomass burning emissions and associated brown carbon formation was directly observed (Gilardoni et al. 2016). HR-ToF-AMS observations have shown

similarity between atmospheric organic matter in fog water and aerosol formed following fog dissipation, indicating low volatility organics that were originally present in the fog are left behind upon evaporation; these particles are enriched in oxidized organic matter, and absorb solar radiation more efficiently than fresh emissions, contributing to atmospheric brown carbon (Gilardoni et al. 2016).

In this study, we analyzed fog from SPC and aerosol from Bologna, collected during the 2013 Supersito field campaign. Due to the intense time investment required for FT-ICR MS data analysis, we chose to focus our detailed analysis on a subset of samples, including two aerosol and two fog samples. The subset was selected to represent the influence of fresh and aged biomass burning emissions on fog and aerosol based on the HR-ToF-AMS and ^1H -NMR observations (see section 2.3). We used a combination of ^1H -NMR, HR-ToF-AMS and FT-ICR MS techniques, to explore the molecular level details of the complex mixtures of atmospheric organic matter in the Po Valley. Similar studies focusing on analysis of atmospheric samples with ^1H -NMR and FT-ICR MS have been conducted in the past (Schmitt-Kopplin et al. 2010, Willoughby et al. 2016), but so far, this type of study with a focus on biomass burning and aqueous phase processing has not been previously reported.

2.3 Methods

2.3.1 Sample collection and chemical analysis

Sub-micrometer (PM_{10}) aerosol particles were collected in Bologna on pre-washed and pre-baked quartz fiber filters (PALL, 18 cm diameter) by a High-Volume Sampler (TECORA Echo Hi Vol) equipped with a digital PM_{10} sampling inlet, at a nominal flow rate of 500 L min^{-1} . PM_{10} samples were collected during winter 2013 (from 4-Feb 2013 to 15-Feb 2013), during the Supersito project. Fog water was collected at the SPC field station, where monitoring of fog occurrence and fog water collection has been performed every year systematically since 1989 (Giulianelli et al. 2014); during the 2013 winter fog samples were collected from 29-Nov 2012 to 12-Mar 2013. In the fog collector (Fuzzi et al. 1997), a short wind tunnel is created by a rear fan, where an air stream containing fog droplets are collected by impaction using a series of stainless steel strings. The collected droplets drain off the strings into a sampling bottle. The air flow through the tunnel was $17 \text{ m}^3 \text{ min}^{-1}$ with a 50% collection efficiency for individual strings ($3 \mu\text{m}$ radius each). All parts of the fog collector coming into contact with the fog droplets, were made of stainless steel to avoid sampling artifacts from adsorption of organic compounds to the surfaces.

The aerosol filters were extracted with deionized ultra-pure water (Milli-Q) in an ultrasonic bath for 1 h. The water extract was filtered with a $0.45 \mu\text{m}$ PTFE membrane in order to remove suspended particles. Fog water was filtered through 47 mm quartz fiber filters within a few hours of collection and conductivity and pH measurements were taken (Crison microCM 2201 conductimeter and Crison micropH 2002 pH meter). Aliquots of both aerosol water extracts and fog water prepared in this way were used to determine the total

organic carbon content (Multi N/C 2100 analyzer; Analytik Jena, Germany) and water soluble organic carbon (WSOC) concentration, (Rinaldi et al. 2007) as well as for ^1H -NMR analysis and HR-ToF-AMS analysis of fog samples described below (HR-ToF-AMS data for aerosol samples was collected in real time).

2.3.2 ^1H -NMR analysis

Aliquots of the aerosol extract and fog water were dried under vacuum and re-dissolved in deuterium oxide (D_2O) for organic functional groups characterization by ^1H -NMR spectroscopy, as described in Decesari et al. (2000). The ^1H -NMR spectra were acquired at 600 MHz (Varian Unity INOVA spectrometer) with a 5 mm probe. Sodium 3-trimethylsilyl-(2,2,3,3- d_4) propionate (TSP- d_4) was used as an internal standard by adding 50 μl of a 0.05% TSP- d_4 (by weight) in D_2O to the standard in the probe. The speciation of hydrogen atoms bound to carbon atoms can be provided by ^1H -NMR spectroscopy in protic solvents, on the basis of the range of frequency shifts, the signal can be attributed to H-C containing specific functionalities (Decesari et al. 2000, Decesari et al. 2007). Detection limits for an average sampling volume of 500 m^3 were of the order of 3 nmol m^{-3} for each functional group. ^1H -NMR spectra were collected during the winter 2013 campaign using the method described above to identify and quantify major components of WSOC in both fog and aerosol. In the present study, the results of these ^1H -NMR analyses were used to characterize and to select the samples for subsequent FT-ICR MS analysis as described in section 2.3.4.

2.3.3 HR-ToF-AMS analysis

During the Supersito winter 2013 campaign (4 February 2013 to 15 February 2013) the chemical composition of PM_{10} aerosol particles at Bologna was characterized with a 5 minute-time resolution using an HR-ToF-AMS (Aerodyne Research (DeCarlo et al. 2006)). Data was collected in the V-ion mode, at a resolution of 2,200. The influx of aerosol particles was dried below 30% relative humidity with a Nafion drier before analysis. Details on analysis of HR-ToF-AMS data for the Supersito winter 2013 campaign were previously reported (Gillardoni et al. 2016); here we report HR-ToF-AMS characterization averaged over the sampling periods of the selected aerosol samples.

Fog water samples were also analyzed by HR-ToF-AMS after being re-aerosolized (TSI constant output atomizer, Model #3076) in an inert argon gas flow, to characterize dissolved water-soluble organics. To make sure that the re-aerosolized fog water represented the original sample, we verified that the nitrate-to-organic carbon and the sulfate-to-organic carbon ratios from the HR-ToF-AMS analysis were within 20% (measurement uncertainty level) of the ratios measured off-line by ion chromatography and thermo-optical analysis.

2.3.4 Ultrahigh resolution FT-ICR MS analysis

Four samples were selected for FT-ICR MS analysis based on the characterization by HR-ToF-AMS data and ^1H -NMR spectra for the entire Supersito winter 2013 sample set (Section 2.3.1). High molecular weight WSOC compounds were prepared for FT-ICR MS analysis using a polymeric reversed phase solid phase extraction (SPE) cartridge (Strata-X, Phenomenex) to remove salts and low molecular weight compounds which interfere with electrospray ionization (ESI). The cartridges were loaded with HCl acidified aqueous samples ($\text{pH} < 2$), rinsed with 1 mL of water, and then eluted using 2 mL of ACN:H₂O (90:10 by volume). Fog samples were later re-filtered using a 25 mm quartz filter before SPE. A portion of the aerosol filter samples were extracted with ultrapure water using sonication and the extracts were then filtered using a 25 mm quartz filter to remove insoluble materials; the aerosol extracts were then prepared for FT-ICR MS analysis using SPE as described above. The WSOC described in this paper is operationally defined as the WSOC that is both retained and recovered from the SPE cartridges (SPE-recovered), thus it is not equivalent with the total WSOC. The ACN:H₂O extracts were analyzed at the Woods Hole Oceanographic Institute in Woods Hole, MA, using full-scan ESI ultrahigh-resolution FT-ICR MS (7T LTQ FT-ICR MS, Thermo Scientific) at a resolving power of 400,000 as described in our previous work (Zhao et al. 2013, Dzepina et al. 2015). We used direct infusion analysis to collect mass spectrometry data over the mass range of m/z 100 to 1000 in the negative ionization mode, for approximately 200 scans. Molecular formulas were assigned as previously described in our work (Mazzoleni et al. 2010, Putman et al. 2012, Zhao et al. 2013, Dzepina et al. 2015) using Sierra Analytics Composer software (version 1.0.5) within the limits of $\text{C}_{2-200}\text{H}_{4-1000}\text{O}_{1-20}\text{N}_{0-3}\text{S}_{0-1}$. The formulas were reviewed manually for their credibility; for further details, see section 2.7. Approximately 74% of the measured masses in each of the samples were assigned a molecular formula. Oxygen to carbon (O:C) and hydrogen to carbon (H:C) ratios, were calculated from the respective number of C, H or O atoms in the assigned molecular formulas. We calculated Kendrick mass (KM) and Kendrick mass defect (KMD) as described in equations (2.1) and (2.2), respectively (Stenson et al. 2003).

$$KM = \text{experimental mass} * \left(\frac{14.00000}{14.01565} \right) \quad (2.1)$$

$$KMD = \text{nominal mass} - KM \quad (2.2)$$

DBE was calculated by equation (2.3) for the molecular formula format: $\text{C}_c\text{H}_h\text{O}_o\text{N}_n\text{S}_s$.

$$DBE = c - \left(\frac{h}{2} \right) + \left(\frac{n}{2} \right) + 1 \quad (2.3)$$

Note that S and O are divalent in equation (2.3); additional unsaturated bonds associated with pentavalent nitrogen, and tetravalent or hexavalent sulfur are not included in this DBE calculation. The average oxidation state of carbon (OS_C) in the molecular formulas was estimated using equation (2.4), based on the approximation described in Kroll et al. (2011); note that the inclusion of nitrogen and sulfur affects the oxidation state of carbon, and

equation (2.4) assumes both are fully oxidized. The modified aromaticity index (AI_{mod}) (Koch and Dittmar 2006, 2016) was calculated using equations (2.5) to (2.7). Equations (2.4) to (2.7) use the same molecular formula format as DBE in equation (2.3).

$$OS_c \approx 2\frac{o}{c} - \frac{h}{c} - 5\frac{n}{c} - 6\frac{s}{c} \quad (2.4)$$

$$DBE_{AI} = 1 + c - \left(\frac{o}{2}\right) - s - \left(\frac{n+h}{2}\right) \quad (2.5)$$

$$C_{AI} = c - \left(\frac{o}{2}\right) - n - s \quad (2.6)$$

$$AI_{mod} = \frac{DBE_{AI}}{C_{AI}} \quad (2.7)$$

In equation (2.7), the $AI_{mod} = 0$, if $DBE_{AI} \leq 0$ or $C_{AI} \leq 0$, as defined in Koch and Dittmar (2006, 2016).

The resulting data set represents the SPE-recovered higher molecular weight water soluble organic aerosol and is expected to predominantly contain acidic compounds due to the negative ion ESI analytical bias. The observed molecular compositions represent the oxidized fraction of the atmospheric samples thus, useful insights can be made with these limitations in mind. Furthermore, it is important to note that the individual molecular formulas likely represent a mixture of structural isomers co-existing in atmospheric organic matter, as recently observed for deep-sea organic matter (Zark et al. 2017).

2.4 Results and discussion

2.4.1 Selection of aerosol and fog water samples

Among the 15 fog and 18 aerosol samples collected during the winter of 2013 at SPC and Bologna, we selected two fog and two aerosol samples for subsequent analysis by FT-ICR MS according to the following rationale. Aerosol samples were selected based on PMF source apportionment of “fresh” and “aged” wood burning emissions using HR-ToF-AMS and 1H -NMR data, as described in Gilardoni et al. (2016). On 13-Feb 2013, a high concentration of SOA was observed, where the ratio of SOA to POA was ~ 4 , and the aqueous SOA from biomass burning accounted for about 55% of total SOA. Thus, BO0213D was defined as strongly influenced by aged wood burning emissions. During the night of 4-Feb 2013, the fresh biomass burning concentration was $\sim 6 \mu g m^{-3}$, accounting for the 54% of total organic aerosol. Thus, BO0204N was defined as strongly influenced by fresh wood burning emissions. Similarly, HR-ToF-AMS observations were used to select fog samples strongly impacted by “fresh” and “aged” wood burning emissions. Specifically, we used the relative intensity of m/z 60 (f_{60}) as a marker of fresh biomass burning influence and m/z 44 (f_{44}) as a marker of oxygenated and processed dissolved organic molecules (Aiken et al. 2008, Gilardoni et al. 2016). The f_{44} vs. f_{60} space was

Figure 2.1. Preliminary characterization of fog and PM₁ aerosol samples collected in SPC and Bologna, respectively, during the 2013 Supersito field campaign. Characterization was performed via high resolution time-of-flight aerosol mass spectrometry (HR-ToF-AMS) analysis as described by Cubison et al. (2011), utilizing the relative intensity of peak m/z 60 (f_{60}) and peak m/z 44 (f_{44}) as markers of fresh biomass burning influence and oxygenated and processed dissolved organic molecules respectively (a). Further characterization was performed via ¹H-NMR analysis, as described by Decesari et al. (2007), where samples were mapped by ¹H-NMR functional group fractions (b). In (b), dashed lines indicate the boundaries of the source fingerprints according to Decesari et al. (2007), (“BB”: biomass burning aerosol) and the x and y axes report the contributions of alkoxy (H-C-O) and acyl (H-C-C=O) groups to the total aliphatic fraction of WSOC respectively. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

previously proposed to represent biomass burning vs. atmospheric aerosol aging (Cubison et al. 2011) and was extended here to fog samples. This representation is an oversimplification of the complexity of organic molecules in fog water, employed here exclusively to note the major differences in terms of emission sources and atmospheric history. In Fig. 2.1a, it can be seen that the fog sample SPC0106F had low f_{44} and high f_{60} values, while SPC0201F had high f_{44} and low f_{60} values. Thus, from here on, SPC0106F (fog) and BO0204N (aerosol) will be referred to as the “fresh” biomass burning influenced samples, and SPC0201F (fog) and BO0213D (aerosol) will be referred to as the “aged” biomass burning influenced samples. A summary of the sample collection details and HR-ToF-AMS characterization is given in Table 2.1.

2.4.2 ^1H -NMR composition

Functional group distributions for the selected PM_{10} and fog samples were provided by ^1H -NMR analysis. A synthetic representation of the ^1H -NMR organic functional groups distribution of all the collected samples is reported in Fig. 2.1b, following the approach described by Decesari et al. (2007) for source attribution. Briefly, Decesari et al. (2007) presented a survey of ^1H -NMR functional group distributions of WSOC samples from diverse environments proposing fingerprints for broad categories of oxygenated organic compounds in aerosol. These categories are: SOA (enriched in acyl groups, H-C-C=O), biomass burning aerosol (enriched in alkoxy, H-C-O , and aromatics), and marine organic aerosol (enriched in aliphatic groups other than acyls and alkoxy, mainly amines and sulfoxy groups). In this study, most samples were categorized either as SOA or biomass burning, even if a significant fraction of the aerosol samples exhibited ^1H -NMR compositions with a very high alkoxy contribution, exceeding the boundaries proposed by Decesari et al. (2007). For example, sample BO0204N (representative of fresh biomass burning aerosol) showed by far the largest contribution of alkoxy groups and the least amount of acyl groups. In contrast, BO0213D (representative of aged aerosols) showed relatively high acyl content and small alkoxy fractions. Similarly, the two selected fog samples (SPC0106F: fresh, and SPC0201F: aged) were clearly differentiated based on their ^1H -NMR functional group distributions (Fig 2.1b). Therefore, the selected aerosol and fog samples represent extremes in the structural space of this WSOC sample set based on the distribution of ^1H -NMR functionalities, and in agreement to the categorization provided by the HR-ToF-AMS measurements.

The differences between the two aerosol samples likely reflect the ambient conditions during sampling: BO0204N was characterized by night-time accumulation of ground-level local emissions from residential heating and an absence of photochemical processes; instead, BO0213D was characterized by daytime photo-chemically processed aerosol and by an enhanced mixing with regional-scale air masses. Similarly, the diversity in the fog samples reflects the collection duration and the associated liquid water content (LWC) of the two considered fog events: SPC0106F was collected over a shorter duration with a lower LWC compared to SPC0201F (Table 2.1).

Table 2.1. Po Valley sample collection, identification and HR-ToF-AMS data.

Sample name	SPC0106F^a	SPC0201F^a	BO0204N^b	BO0213D^b
Collection site	San Pietro Capofiume	San Pietro Capofiume	Bologna	Bologna
Sample type	Fog water	Fog water	PM ₁ aerosol	PM ₁ aerosol
Fresh vs. Aged influence	Fresh	Aged	Fresh	Aged
Start collection date and time^c	6-Jan 2013 3:10	1-Feb 2013 19:40	4-Feb 2013 18:18	13-Feb 2013 9:24
Collection time (h)	1.33	15.37	14.62	8.60
Temperature (°C)^b	1.0	3.0	5.9	3.0
pH	5.81	3.34	NA	NA
[NO_x] (ppb)^d	73	15	146	101
Relative humidity (%)^{d, e}	100	100	58	80
liquid water content (mL m⁻³)^d	0.190	0.258	NA	NA
aerosol liquid water content (μg m⁻³)^{d, f}	NA	NA	69	515
f₄₄^g	0.16	0.21	0.042 ± 0.006	0.097 ± 0.004
f₆₀^g	0.007	0.004	0.016 ± 0.003	0.010 ± 0.001
OM:OC^{d, h}	1.9	2.2	1.5 ± 0.1	1.9 ± 0.1
O:C^{d, h}	0.58	0.8	0.24 ± 0.04	0.56 ± 0.03
H:C^{d, h}	1.37	1.29	1.65 ± 0.03	1.60 ± 0.01
OS_C^{d, h}	-0.21	0.32	-1.17 ± 0.08	-0.48 ± 0.06

^aFog water samples were re-aerosolized for HR-ToF-AMS data analysis; ^bAerosol sample data is from on-line measurements and the standard deviation of on-line measurements corresponding to the sample collection period are shown; ^cStart collection times given in local time; ^dAverage values corresponding to the collection times of individual samples; ^eAverage RH was assumed to be 100% for fog samples, as super-saturation levels could not be measured; ^fALWC is an average of E-AIM and ISORROPIA modeled data for the sampling period; ^gFractional abundance of a mass fragment (f_x) was calculated as the ratio between that fragment signal and the total organic concentration; ^hElemental ratios were calculated according to Aiken et al. (2008).

Figure 2.2. The ^1H -NMR spectra of selected fog water (a and b) and aerosol (c and d) samples, and their corresponding functional groups distribution. A set of specific resonances was attributed to individual compounds: 1) formate, 2) phthalic acid, 3) ammonium, 4) maleic acid, 5) levoglucosan, 6) hydroxy-methanesulfonic acid, 7) trimethylamine, 8) methanesulfonic acid, 9) dimethylamine, 10) monomethylamine, 11) succinic acid, 12) pyruvic acid, 13) lactic acid. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

It should be noted that although a pair of fresh and aged samples were selected from each of the sample sets, Fig. 2.1 shows a clear shift in the average composition between the fog and the aerosol samples, where the fog samples were characterized by a greater amount of acyl groups and a smaller fraction of alkoxylys. So, according to the simple source-attribution scheme based on the major ^1H -NMR functionalities presented here, the fog compositions were more oxidized and “SOA-like” than aerosols. As a consequence, the fresh fog composition overlapped with the aged aerosol composition (Fig. 2.1b). This implies that the fresh fog sample SPC0106F was processed to a similar degree as the most aged aerosol sample BO0213D. This was confirmed by the corresponding HR-ToF-AMS elemental ratios (very similar O:C for SPC0106F and BO0213D, see Table 2.1) and by the detailed comparison between the ^1H -NMR spectra of these two samples (one fog and one aerosol). This difference in the average functional group composition between fog and aerosol samples in the Po Valley can be explained by: (a) the preferential scavenging of more oxidized constituents of organic particles into fog (Gilardoni et al. 2014); (b) the effect of oxidative chemical reactions in fog water leading to the production of carboxylic acids and carbonyls (hence, acyls); and (c) a stronger aging effect from fog processing at the rural site (SPC) with respect to urban areas (Bologna) at the margins of the Po basin.

The ^1H -NMR spectra of the selected samples are reported in Fig. 2.2. The spectra of the aerosol samples (Fig. 2.2c and 2.2d) exhibited a clear biomass burning fingerprint, with evident proton resonances from levoglucosan and intense bands from alkoxy (H-C-O) and aromatic (Ar-H) groups. However, the band of phenols and methoxyphenols, which are primary biomass burning tracers, were clearly found only in the spectrum of BO0204N, representative of fresh primary organic aerosols in our study. Moreover, the fraction of levoglucosan and alkoxy groups was much greater in BO0204N than in BO0213D. The aged aerosol BO0213D contained higher amounts of two methylamines (mono- and trimethylamines) relative to BO0204N, and especially much larger fractions of methanesulfonate and succinic acid, which are tracers of SOA. The spectral region between the chemical shift of 2.1 and 2.4 ppm showed clear bands representing aliphatic dicarboxylic acids and ketoacids (Suzuki et al. 2001) in the aged aerosol, but were barely visible in the fresh aerosol. The aged aerosol was also characterized by the occurrence of hydroxy-methanesulfonic acid (HMSA), a known tracer of aqueous SOA. Similarly for fog, SPC0106F (Fig. 2.2a) exhibited a clear biomass burning fingerprint with contributions from levoglucosan, alkoxy (H-C-O), and aromatic groups (Ar-H), whereas SPC0201F (Fig. 2.2b) showed tracers of aqueous-phase SOA (HMSA) and high concentrations of acyl groups, (CH-C=O), which demonstrated the effects of the aging process. Additionally, SPC0201F exhibited several low-molecular weight organic acids (phthalic, maleic, succinic, pyruvic and lactic acids) in much greater amounts than SPC0106F, where only traces of phthalic and succinic acids were found. This indicated that the aged fog was enriched in products of the oxidative degradation of particulate and gaseous organic compounds. It should be noted, that the fresh fog (SPC0106F) did not show the prominent band from phenols or methoxyphenols observed in the spectrum of the fresh aerosol (BO0204N). This suggests that the WSOC of the fresh fog had undergone a certain degree of chemical modification respective to primary biomass burning OA.

2.4.3 Ultrahigh resolution FT-ICR MS composition

2.4.3.1 Overview of the Po Valley ambient fog and aerosol compositions

Approximately 1600 to 2800 individual monoisotopic molecular formulas were assigned to the ultrahigh resolution mass spectra of the SPE-recovered WSOC from each Po Valley sample. Based on the inclusion of C, H, N, O, and S elements, the molecular formulas were sorted into the following elemental groups: “CHO,” “CHNO,” “CHOS” and “CHNOS.” The percent composition of these elemental groups for each sample is shown in Fig. 2.3. Most of the molecular formulas were present in the subclasses O_{4-10} , NO_{3-13} , $O_{5-10}S$ and $NO_{7-11}S$ (Fig. 2.4). A summary of the observed numbers of formulas per elemental group, as well as the average O:C, H:C, OS_C and DBE values are provided in Table 2.2. Although they are not expected to match, the values for the SPE-recovered WSOC do trend with those from the HR-ToF-AMS data shown in Table 2.1; we note, that not only are the elemental ratios from different fractions of the aerosol, but they are also determined differently.

A great diversity of CHNO, CHOS and CHNOS formulas were observed in the Po Valley samples, likely representing organonitrates, organosulfates, and nitrooxy-organosulfates. These compound classes can be inferred from the analytical bias of the negative mode ESI, as well as the O:N and O:S of the assigned molecular formulas. Nearly all N-containing formulas had $O:N > 3$, suggesting that a majority of the nitrogen species contained at least one nitro or nitrate group. Multiple nitrogen species, such as those of classes N_2O_{3-5} and N_3O_{5-7} have an O:N low enough to indicate amine, imine, or imidazole structures, as these types of products have been reported in cloud water mimic reactions (De Haan et al. 2011), however only a modest number of formulas with multiple nitrogen atoms were observed. All of the S-containing formulas had $O:S > 4$ ratios, suggesting sulfite, sulfate, and sulfonic acid functionalities. These inferences are consistent with the ionization polarity, where oxidized and acidic components are more efficiently ionized in negative ion ESI. A study by LeClair et al. (2012) who performed FT-ICR MS² using negative mode ESI on a variety of CHNO, CHOS, and CHNOS components confirmed that the studied compounds in Fresno fog were indeed multifunctional organonitrates, organosulfates, and nitrooxy-organosulfates. Furthermore, nitrate and sulfate salts are common secondary components present in the Po Valley (Giulianelli et al. 2014) and reactions between these inorganic salts and organics are expected as secondary reactions in the aqueous phase (Noziere et al. 2010, McNeill et al. 2012, Herrmann et al. 2015, McNeill 2015). Amines have been observed in the Po Valley, emitted by livestock farming and waste treatment activity, and it is possible that some species with amine groups were emitted from smoldering biomass combustion (Andreae and Merlet 2001, Paglione et al. 2014). However, given the analytical bias for acidic functional groups in the ESI negative ion mode, it is unlikely that reduced nitrogen species were detected.

Figure 2.3. Percentage of assigned molecular formulas to each of the elemental groups in the Po Valley samples, where (a) includes all identified molecular formulas and (b) includes only the unique molecular formulas. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

Table 2.2. Summary of Po Valley sample FT-ICR MS formula assignment data.

	All	CHO	CHNO	CHOS	CHNOS
Number	2824	1158 (41%)	744 (26%)	619 (22%)	303 (11%)
Molecular weight (Da)^a	368.44 ± 94.21	359.05 ± 101.35	342.88 ± 90.69	404.68 ± 83.64	393.12 ± 61.72
O:C^a	0.479 ± 0.16	0.415 ± 0.13	0.503 ± 0.14	0.488 ± 0.14	0.642 ± 0.18
H:C^a	1.30 ± 0.36	1.21 ± 0.32	1.14 ± 0.27	1.56 ± 0.34	1.53 ± 0.34
OS_C^a	-0.623 ± 0.48	-0.379 ± 0.42	-0.563 ± 0.33	-0.950 ± 0.49	-1.039 ± 0.37
DBE^a	7.24 ± 3.65	8.29 ± 3.67	8.35 ± 3.01	4.93 ± 3.08	5.20 ± 2.93
AI_{mod}^a	0.24 ± 0.22	0.31 ± 0.20	0.32 ± 0.21	0.08 ± 0.12	0.06 ± 0.11
C_N^a	17.2 ± 5.2	18.3 ± 5.5	15.7 ± 4.9	17.8 ± 4.8	15 ± 3.7
O_N^a	7.8 ± 2.3	7.4 ± 2.5	7.6 ± 2.2	8.3 ± 2.0	9.1 ± 1.3
Number	1671	890 (53%)	427 (26%)	212 (13%)	142 (8%)
Molecular weight (Da)^a	360.12 ± 97.52	358.18 ± 108.61	364.33 ± 90.94	360.12 ± 78.27	359.66 ± 63.56
O:C^a	0.577 ± 0.18	0.509 ± 0.13	0.617 ± 0.14	0.592 ± 0.15	0.858 ± 0.24
H:C^a	1.31 ± 0.35	1.18 ± 0.29	1.23 ± 0.26	1.74 ± 0.23	1.77 ± 0.20
OS_C^a	-0.399 ± 0.51	-0.161 ± 0.41	-0.369 ± 0.28	-1.006 ± 0.42	-1.075 ± 0.35
DBE^a	6.83 ± 3.53	8.05 ± 3.38	7.49 ± 2.69	3.01 ± 1.81	2.93 ± 1.13
AI_{mod}^a	0.22 ± 0.21	0.30 ± 0.19	0.22 ± 0.20	0.00 ± 0.03	0.00 ± 0.00
C_N^a	15.8 ± 5.0	16.9 ± 5.2	15.4 ± 4.4	14.4 ± 4.0	11.8 ± 3.4
O_N^a	8.7 ± 2.7	8.5 ± 3.0	9.2 ± 2.4	8.2 ± 2.0	9.5 ± 1.6

^aMolecular weight, O:C, H:C, oxidation state of carbon (OS_C), double bond equivalents (DBE), modified aromaticity index (AI_{mod}), carbon number (C_N), and oxygen number (O_N) values represent mathematical averages based on formula assignment, with standard deviation provided. These values were obtained using equations (2.1) to (2.7).

(Table 2.2 continued next page)

(Table 2.2 continued from previous page)

	All	CHO	CHNO	CHOS	CHNOS	
BO0204N	Number	1634	808 (49%)	732 (45%)	42 (3%)	52 (3%)
	Molecular weight (Da) ^a	364.99 ± 100.13	358.24 ± 105.13	373.63 ± 98.27	332.45 ± 54.27	374.39 ± 50.93
	O:C ^a	0.433 ± 0.14	0.377 ± 0.11	0.480 ± 0.14	0.405 ± 0.11	0.652 ± 0.17
	H:C ^a	1.13 ± 0.32	1.12 ± 0.30	1.04 ± 0.22	1.96 ± 0.10	1.77 ± 0.14
	OS _C ^a	-0.471 ± 0.41	-0.368 ± 0.38	-0.461 ± 0.28	-1.578 ± 0.22	-1.303 ± 0.21
	DBE ^a	9.26 ± 3.94	9.42 ± 4.03	9.99 ± 3.11	1.31 ± 0.64	3.08 ± 0.9
	AI _{mod} ^a	0.36 ± 0.20	0.38 ± 0.19	0.39 ± 0.18	0.00 ± 0.00	0.00 ± 0.00
	C _N ^a	18.0 ± 5.6	18.9 ± 5.7	17.6 ± 5.5	15.0 ± 3.5	13.9 ± 3.4
	O _N ^a	7.5 ± 2.4	7.0 ± 2.5	8.1 ± 2.3	5.8 ± 0.9	8.6 ± 0.7
BO0213D	Number	2753	1097 (40%)	1123 (41%)	249 (9%)	284 (10%)
	Molecular weight (Da) ^a	361.82 ± 96.19	351.26 ± 102.62	360.99 ± 94.43	354.82 ± 75.64	412.02 ± 76.3
	O:C ^a	0.498 ± 0.19	0.424 ± 0.15	0.555 ± 0.18	0.435 ± 0.16	0.617 ± 0.21
	H:C ^a	1.37 ± 0.37	1.25 ± 0.34	1.26 ± 0.27	1.9 ± 0.22	1.8 ± 0.18
	OS _C ^a	-0.683 ± 0.53	-0.399 ± 0.46	-0.631 ± 0.34	-1.445 ± 0.37	-1.322 ± 0.30
	DBE ^a	6.64 ± 3.65	7.81 ± 3.88	7.44 ± 2.6	1.8 ± 1.43	3.19 ± 1.47
	AI _{mod} ^a	0.21 ± 0.21	0.29 ± 0.21	0.22 ± 0.20	0.01 ± 0.06	0.00 ± 0.00
	C _N ^a	16.7 ± 5.3	17.9 ± 5.7	15.8 ± 4.9	16 ± 4.9	16 ± 4.7
	O _N ^a	7.8 ± 2.5	7.2 ± 2.4	8.3 ± 2.5	6.3 ± 1.4	9.1 ± 1.6

^aMolecular weight, O:C, H:C, oxidation state of carbon (OS_C), double bond equivalents (DBE), modified aromaticity index (AI_{mod}), carbon number (C_N), and oxygen number (O_N) values represent mathematical averages based on formula assignment, with standard deviation provided. These values were obtained using equations (2.1) to (2.7).

Figure 2.4. Distributions of the molecular formulas within all 64 elemental group subclasses for CHO, CHNO, CHOS and CHNOS groups as indicated in the Figure. The total number of molecular formulas for each SPE-recovered WSOC sample were split into two groups of unique and non-unique formulas; the darker shade represents formulas unique to a sample, (denoted in the Figure legend with an asterisk after the sample name, e.g. “Fresh Fog*”) while the lighter shade represents common formulas. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

Figure 2.5. The modified aromaticity index (AI_{mod}) for the assigned molecular formulas and the percentage of each AI_{mod} type, as defined by Koch and Dittmar (2006, 2016): aliphatic ($AI_{\text{mod}} = 0$), olefinic ($0 < AI_{\text{mod}} \leq 0.5$), aromatic ($AI_{\text{mod}} > 0.5$), and condensed aromatic ($AI_{\text{mod}} \geq 0.67$). Here aromatic and condensed aromatic formulas were combined, because a small fraction of condensed aromatics was observed. The results are partitioned by elemental group, where it can be seen that the majority of olefinic and aromatic compounds belong to the CHO and CHNO groups. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

Nitrated phenols are known contributors to light absorbing atmospheric brown carbon and are associated with biomass burning (Desyaterik et al. 2013, Laskin et al. 2015). In this work, a large number of CHNO formulas were observed with low H:C and low O:C, especially in the fresh aerosol and fresh fog samples; several of the CHNO formulas were also estimated to be aromatic using the AI_{mod} calculation (Fig. 2.5). Specifically, the molecular formulas for nitrophenol ($C_6H_5NO_3$), methyl-nitrophenol ($C_7H_7NO_3$), dinitrophenol ($C_6H_4N_2O_5$), nitroguaiacol ($C_7H_7NO_4$), and nitrosalicylic acid ($C_7H_5NO_5$) (Kitanovski et al. 2012, Desyaterik et al. 2013) were observed in all four Po Valley samples.

All of the molecular formulas were plotted in van Krevelen space (H:C vs. O:C) partitioned by sample (columns) and elemental group (rows) (Fig. 2.6). In this space, molecular formulas with $O:C \geq 0.6$ and $OS_C \geq 0$ are considered to be highly oxidized and formulas with $H:C \geq 1.2$ are considered to be highly saturated (Tu et al. 2016). The distribution of the CHO and CHNO formulas is quite similar to WSOC extracted from ambient fog collected in Fresno, CA (Mazzoleni et al. 2010). Additionally, the distribution of CHO formulas from phenolic aqueous SOA reported in Yu et al. (2016) partially covers the same area of the van Krevelen space. The CHOS and CHNOS formulas with high H:C ratios were also distributed similarly to Mazzoleni et al. (2010). The high H:C ratios indicate that a majority of the CHOS and CHNOS formulas represent aliphatic organosulfate compounds, consistent with the aliphatic AI_{mod} values (Fig. 2.5). In contrast, a majority of the formulas with aromatic AI_{mod} values were in the CHO and CHNO groups, and tended to cluster at low H:C and low O:C in the van Krevelen space, in agreement with previous studies (Mazzoleni et al. 2010, LeClair et al. 2012). Consistent with the 1H -NMR results in Fig. 2.1b, the van Krevelen diagrams for SPC0106F and BO0213D were similar (see also Fig. 2.6), barring the additional low H:C CHOS and CHNOS formulas of SPC0106F and the additional CHNO formulas of BO0213D.

Underscoring the influence of biomass burning on these samples, we found several molecular formulas matching previously observed species in biomass burning influenced ambient cloud water from Mt. Tai, China (Desyaterik et al. 2013) (Appendix A: Table 6.1). There were also several matches with the products of laboratory phenolic aqueous SOA reactions (Yu et al. 2014, Yu et al. 2016) (Appendix A: Table 6.2). Other notable molecular formulas included those for the compounds: acetosyringone, acetovanillone, azelaic acid, benzoic acid, coumaric acid, hydroxybenzoic acid, ketolimononaldehyde, nitrocatechol, o-toluic acid, phthalic acid, syringaldehyde, syringic acid, tyrosine, vanillic acid and vanillin (Mazzoleni et al. 2007, Desyaterik et al. 2013, Nguyen et al. 2013, Pietrogrande et al. 2014, Pietrogrande et al. 2014, Yu et al. 2014, Dzepina et al. 2015, Pietrogrande et al. 2015, Yu et al. 2016). The molecular formulas for common methoxyphenols, syringol ($C_8H_{10}O_3$), methylsyringol ($C_9H_{12}O_3$), and eugenol ($C_{10}H_{12}O_2$), were present in all samples except BO0204N; as they are both semi-volatile and water-soluble, they are not expected to be present in aerosol with low liquid water content.

Figure 2.6. van Krevelen diagrams for the SPE-recovered water soluble organic carbon by elemental group (rows) and sample (columns) as indicated in the Figure. Dashed lines represent $H:C = 1.2$ (horizontal), $O:C = 0.6$ (vertical) and $OSC = 0$ (diagonal) as described in Tu et al. (2016). Formulas unique to each sample are color scaled to the number of oxygen atoms in the assigned formula; grey points represent common molecular formula assignments. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

Several formulas were also found that could be more oxidized versions of phenolic species produced from biomass burning. These formulas included additional oxygen atoms added to the base formulas for phenol ($C_6H_6O_{3-5}$), guaiacol ($C_7H_8O_{3-6}$) and syringol ($C_8H_{10}O_{4-7}$). Five of these formulas, $C_6H_6O_3$, $C_6H_6O_5$, $C_8H_{10}O_5$, $C_8H_{10}O_6$ and $C_8H_{10}O_7$, were previously observed in biomass burning aerosol (Pietrogrande et al. 2015) and in the products of laboratory phenolic aqueous phase SOA reactions (Yu et al. 2014, Yu et al. 2016).

2.4.3.2 Molecular trends for ambient fog and aerosol compositions

Molecular formula trends in the form of histograms are a useful way to organize and visualize the thousands of formulas observed here. The trends based on carbon number, oxygen number, and DBE of the assigned molecular formulas are shown in Fig. 2.7. Although relative abundance does not directly correspond to analyte concentrations, it provides a basis for relative comparisons. For example, the influence of terpene SOA products is indicated from the elevated total relative abundance of molecular formulas near C_{10} (observed in all samples) and an additional increased abundance between C_{15-18} (observed in most samples). This was especially pronounced in BO0213D (Fig. 2.7a). These formulas are likely derived from monoterpenes (C_{10}) and sesquiterpenes (C_{15}), where terpene emissions have been observed in biomass burning (Andreae and Merlet 2001). Terpene oxidation products, including organosulfates, were previously observed in biomass burning influenced cloud water (Zhao et al. 2013, Cook et al. 2017) and many of the same molecular formulas were observed in this study (Appendix A: Table 6.3). Specifically, we observed molecular formulas for pinic acid ($C_9H_{14}O_4$), ketopinic acid ($C_{10}H_{14}O_3$), pinonic acid ($C_{10}H_{16}O_3$), hydroxy-dimethylglutaric acid ($C_7H_{12}O_5$), and methyl-butanetricarboxylic acid ($C_8H_{12}O_6$) (He et al. 2014). Overall, the trends indicate an enhanced abundance of CHO formulas in SPC0201F, CHNO formulas in BO0204N and BO0213D, and CHOS formulas in SPC0106F (Fig 2.7a). Consistent with the 1H -NMR results in Fig. 2.1b, there is a strong similarity between samples SPC0106F and BO0213D, especially for the oxygen and DBE trends shown in Figs. 2.7b and 2.7c.

Difference mass spectra were constructed from the assigned monoisotopic molecular formulas for the fog and aerosol samples (Fig. 2.8) and provide a direct comparison of their compositions. Each of the individual relative abundances were normalized by the total abundance of the assigned masses for each sample. In Fig. 2.8, the individual masses with higher abundances in either the positive or negative direction were substantially greater in the fresh or aged samples, respectively; the masses with similar relative abundances tended to cancel each other. Overall, we observed molecular formulas with higher oxygen content at lower molecular weights in the two aged samples, compared to the two fresh samples. To investigate this further, we adapted the approach used for the molecular formula trends described above with the difference relative abundances. The resulting difference trend plots are shown in Fig. 2.9 for carbon, and Figs. 2.10 and 2.11 for oxygen and DBE respectively. In Fig. 2.9b, it is clear there was an enhanced abundance of CHOS and CHNOS formulas with higher carbon numbers in the fresh fog, while the aged fog showed an enhanced abundance of low carbon number CHO formulas. In Fig. 2.9a, it is clear that the fresh aerosol had an enhanced abundance of higher carbon number formulas, though

Figure 2.7. Molecular formula trends for carbon (a), oxygen (b) and the number of double bond equivalents (c). All detected molecular formula abundances were normalized to the total assigned ion abundance for each sample and then summed across the integer values for carbon number, oxygen number, or double bond equivalent values. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

Figure 2.8. Reconstructed difference mass spectra for theoretical masses of assigned molecular formulas in the Po Valley samples with normalized relative abundance. Fresh influenced samples (SPC0106F and BO0204N) are plotted with positive abundance and aged influenced samples (SPC0201F and BO0213D) are plotted with negative abundance. Molecular compositions in both samples with the same mass and similar normalized relative abundance are reduced toward zero. The peaks in the mass spectra are color scaled to the number of oxygen atoms in the assigned molecular formula, where it can be observed that the aged samples shift towards species with higher oxygen numbers at lower masses, compared to the fresh samples. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

unlike the fog samples, they were mainly CHO and CHNO compounds. The aged aerosol had an enhanced abundance of low carbon number formulas from the CHOS and CHNOS groups. In both fog and aerosol, there is an enhanced abundance of higher carbon numbers in the fresh samples relative to the aged samples. Overall, the carbon numbers are shifted to lower values in the fog compared to aerosol (Fig. 2.9) and the oxygen numbers are shifted to higher values in fog compared to aerosol (Fig. 2.10).

The subsequent sections discuss the molecular diversity of the different samples, especially considering the sample type, atmospheric processes during sample collection and the unique molecular formulas observed in each sample. The distributions of unique molecular formulas are shown in Fig. 2.3b and Fig. 2.4.

2.4.3.3 Comparison of the fresh and aged biomass burning influenced fog compositions

The molecular formulas of the aged biomass burning influenced fog (SPC0201F) were more oxidized than the fresh biomass burning influenced fog (SPC0106F). This enhancement in oxidation is shown in Fig. 2.10a, with a greater abundance of higher oxygen number formulas observed in the aged fog. The opposite is true for DBE and carbon numbers, where both trended to higher numbers in the fresh fog compared to the aged fog (See Figs. 2.9b and 2.11b). Most of the CHOS and CHNOS formulas in SPC0106F and SPC0201F were classified as aliphatic by AI_{mod} and approximately 30% of these formulas in SPC0106F were classified as olefinic, which was higher than any other sample (Fig. 2.5). This suggests that the fresh fog molecular formulas represented molecules with large unsaturated carbon backbones, which is consistent with pollutants without significant atmospheric aging. In contrast, the molecular formulas that were more oxidized with smaller carbon backbones were more prevalent in the aged fog.

The unique molecular formulas found in the fresh fog (SPC0106F) were mostly of the $O_{5-13}S$ and $NO_{7-12}S$ subclasses. Organosulfates are known products of aqueous secondary processes, (Darer et al. 2011, Ervens et al. 2011, Schindelka et al. 2013, McNeill 2015) and nucleation scavenging from the preceding fog nuclei composition likely plays a significant role as well (Herckes et al. 2007, Darer et al. 2011, Hu et al. 2011, Gilardoni et al. 2014). The aromatic organosulfates and nitrooxy-organosulfates observed in fresh biomass burning aerosol (Staudt et al. 2014) were not observed here. In general, organosulfates are the products of aqueous-phase SOA reactions which are expected to be enhanced at acidic pH (Noziere et al. 2010, Ervens et al. 2011, McNeill et al. 2012). Because the pH of SPC0106F was only slightly acidic at 5.81, we propose that the formation of these organosulfates may have been promoted by low LWC, and thus relatively high solute concentrations, during the activation of the fog droplets or possibly in the fully formed fog droplets. Organosulfates may also efficiently nucleate droplets, leading to their eventual presence in the fog samples. A noticeable number of CHOS and CHNOS formulas unique to SPC0106F had higher DBE values than formulas from other samples. There was an overall preference for CHOS and CHNOS formulas with DBE values < 6 , except for some of the formulas in SPC0106F which were higher (Fig. 2.7c).

Figure 2.9. Carbon difference trend plots for aerosol (a) and fog (b) sample types. Difference trends were calculated as the difference in abundance for each molecular formula between the respective aged and fresh samples, then summed for each integer carbon number value. Positive values indicate an enhanced relative abundance of the formulas in the fresh sample compared to the aged sample. Similarly, negative values indicate an enhanced abundance of formulas in the aged sample compared to the fresh sample. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

Figure 2.10. Oxygen difference trend plots for aerosol (a) and fog (b) sample types. Difference trends were calculated as the difference in abundance for each molecular formula between the respective aged and fresh samples, then summed for each integer oxygen number value. Positive values indicate an enhanced relative abundance of the formulas in the fresh sample compared to the aged sample. Similarly, negative values indicate an enhanced abundance of formulas in the aged sample compared to the fresh sample. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

Figure 2.11. Double bond equivalent difference trends for aerosol (a) and fog (b) sample types. Difference trends were calculated as the difference in abundance for each molecular formula between the respective aged and fresh samples, then summed for each integer double bond equivalent value. Positive values indicate an enhanced relative abundance of the formulas in the fresh sample compared to the aged sample. Similarly, negative values indicate an enhanced abundance of formulas in the aged sample compared to the fresh sample. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

The 10 most abundant unique molecular formulas in the fresh biomass burning influenced fog of SPC0106F were all CHOS and CHNOS formulas: C₁₁H₁₅NO₈S, C₁₃H₁₄O₈S, C₁₄H₁₆O₈S, C₁₅H₁₆O₉S, C₁₅H₂₄O₇S, C₁₅H₂₄O₈S, C₁₆H₁₈O₉S, C₁₇H₂₀O₉S, C₁₈H₃₀O₈S and C₁₉H₂₄O₉S. These formulas may be tracer species for partially fog processed biomass burning emissions.

While all samples contained some unique molecular formulas among the CHO subclasses, a high number of formulas in the O₉₋₁₄ subclasses were unique to the aged fog (SPC0201F). This trend could indicate enhanced oxidation and aging as a result of aqueous phase reactions in fog. The high average O:C ratio (0.577 ± 0.18) and low pH (3.34) of SPC0201F is consistent with the trend observed by Cook et al. (2017) for cloud water, where the average O:C increased with decreasing pH. Overall, we observed a significant number of CHNO, CHOS and CHNOS molecular formulas in SPC0201F, which are expected products of secondary aqueous phase reactions in fog. However, there was a lower percentage of CHNO, CHOS and CHNOS formulas, and an increased percentage of CHO formulas in SPC0201F compared to SPC0106F, suggesting that aqueous SOA products with N or S may have been transformed by acid hydrolysis into more stable CHO species (Darer et al. 2011). This is reasonable given the longer duration of the fog episode as well as the higher LWC of SPC0201F compared to SPC0106F. The increased oxidation is supported by the ¹H-NMR analysis, which showed an enrichment of carboxylic acids and other compounds carrying acyl groups. SPC0201F had additional unique formulas which were highly oxygenated in the NO₁₃, O₁₁S and NO₁₃S subclasses, which appeared on the low mass end of the homologous series in the CHOS and CHNOS groups (Fig. 2.12). The 10 most abundant unique molecular formulas in the aged biomass burning influenced fog (SPC0201F) were CHO, CHOS and CHNOS species with smaller carbon skeletons than the fresh biomass burning influenced fog (SPC0106F), including: C₄H₉NO₇S, C₅H₉NO₇S, C₈H₁₂O₇S, C₈H₁₃NO₁₁S, C₈H₁₄O₇S, C₉H₁₆O₈S, C₁₀H₁₀O₇, C₁₀H₁₈O₅S, C₁₁H₈O₇ and C₁₂H₁₄O₉. These formulas may be tracer species for heavily fog processed biomass burning emissions.

2.4.3.4 Comparison of the fresh and aged biomass burning influenced aerosol compositions

Similar to the fog samples, the fresh aerosol formulas trended towards higher carbon and DBE numbers relative to the aged aerosol formulas. These carbon number and DBE trends are clearly visible through the difference trends shown in Figs. 2.9a and 2.11a, respectively. Both aerosol samples had a high percentage of formulas that contained nitrogen, with a noticeable number of CHNO formulas unique to these samples (Fig. 2.3b). This larger percentage of CHNO formulas may be attributed enhanced NO_x concentrations associated with urban traffic emissions (Glasius et al. 2006). However, residential wood combustion influenced cloud water collected near Steamboat Springs, CO, was found to be composed of ~52% CHNO molecular formulas (Zhao et al. 2013), and elevated numbers of CHNO formulas were also reported in aerosol with a strong regional biomass burning influence (Schmitt-Kopplin et al. 2010) and wildfire influenced cloud water (Cook et al. 2017).

Figure 2.12. Kendrick mass defect diagrams for each of the Po Valley samples, partitioned by elemental group (rows) and sample (columns) as indicated in the Figure. The molecular formulas unique to each sample are color scaled to the number of oxygen atoms in the assigned formula; grey points represent formulas which are common. Homologous series of molecular formulas are visible as horizontal rows of points, where formulas which are unique to a sample may make up all or only part of an individual homologous series. The sample names Fresh Fog, Aged Fog, Fresh Aerosol, and Aged Aerosol correspond to SPC0106F, SPC0201F, BO0204N, and BO0213D, respectively.

A majority of the unique formulas in the fresh aerosol (BO0204N) were in the NO₆₋₁₂ and N₂O₇₋₁₁ subclasses, which were expected to be products of NO_x reactions and nighttime nitrate radical reactions. The formulas of BO0204N were less saturated and less oxygenated, compared to the formulas in the aged aerosol (BO0213D), which would be expected with little to no influence from aqueous phase secondary processes in the dry conditions of BO0204N (Ervens et al. 2011, McNeill 2015). This could also help to explain the low percentage of CHOS and CHNOS formulas observed in BO0204N. Overall, no unique CHOS formulas were detected in BO0204N and only 1 unique CHNOS formula (C₈H₁₁NO₈S) was detected. The small number of observed CHOS and CHNOS formulas in BO0204N may have originated from the increase in LWC (up to ~300 µg m⁻³) observed in the last 4 hours of sample collection, and thus may have been formed by similar processes as in BO0213D. DBE values in BO0204N trended towards values up to 10, which was much higher than in other samples, where the trend stopped near the DBE of 5 (Fig. 2.7c). The 10 most abundant unique molecular formulas in the aerosol with a fresh biomass burning influence (BO0204N) were mostly N₁ and N₂ CHNO formulas: C₈H₄N₂O₆, C₁₂H₁₀N₂O₈, C₁₃H₁₂N₂O₈, C₁₅H₁₄N₂O₁₀, C₁₆H₁₅NO₆, C₁₇H₂₀O₅, C₂₀H₁₈O₈, C₂₄H₂₁NO₁₀, C₂₄H₂₃NO₁₀ and C₂₆H₂₃NO₁₀. These formulas may be tracer species for biomass burning emissions when nighttime gas phase reactions are dominant.

Several unique molecular formulas for the aged aerosol (BO0213D) were found in the N₂O₄₋₁₃ and N₃O₅₋₁₃ subclasses, as well as the O₄₋₇S, NO₅₋₇S and NO₁₀₋₁₂S subclasses. A large fraction of the N₂ formulas, and all of the N₃ formulas were unique to BO0213D. Compared to the other samples, BO0213D was collected during relatively high NO_x conditions, as well as high humidity and aerosol liquid water content compared to the other aerosol sample. The increased frequency of CHOS and CHNOS formulas in BO0213D compared to BO0204N was likely from reactions in the aqueous phase, enhanced by the increased concentration of species in aerosol liquid water (Darer et al. 2011, Hu et al. 2011, McNeill et al. 2012). Accretion reactions such as aldol condensation, acetal, and hemiacetal reactions are also expected to take place at a significant rate in these enhanced concentrations (Herrmann et al. 2015). While there was not a significant trend towards higher masses in BO0213D compared to other samples, the unique molecular formulas of this sample tended to fall on the high mass end of the homologous series, especially for CHNOS formulas (Fig. 2.12). The 10 most abundant unique molecular formulas for BO0213D were mostly highly oxygenated CHNO formulas: C₇H₉NO₃, C₉H₁₅NO₁₀, C₁₂H₂₅NO₈S, C₁₅H₂₄O₁₂, C₁₆H₁₈N₂O₁₁, C₁₆H₂₀N₂O₁₁, C₁₇H₂₂N₂O₁₁, C₁₇H₂₂N₂O₁₃, C₁₈H₂₁N₃O₁₁ and C₁₈H₂₄N₂O₁₁. These formulas may be tracer species for biomass burning emissions heavily aged by reactions in aerosol liquid water with photolysis.

2.5 Summary and implications

Hygroscopic species are expected to enhance droplet formation, indicating that organics acting as fog nuclei must be somewhat aged. In fog or wet aerosol, the water-soluble organics are subjected to further transformation in the aqueous phase, as we have observed here. These transformation processes in fog and aerosol water were shown to produce

oxygenated and oxidized molecular formulas, as well as N-containing and S-containing formulas with what were likely nitrate and sulfate functional groups. On the basis of the analysis of the selected aerosol and fog samples, representing extreme cases in the HR-ToF-AMS and ^1H -NMR projections of the organic aerosol structural space, we can summarize the following observations:

- An overall molecular trend was observed for both fog and aerosol samples, of concurrent shifts from lower H:C and O:C in samples with fresh biomass burning influence, and toward higher H:C and O:C values in samples with aged biomass burning influence. This was consistent with the ^1H -NMR functional group distributions, which showed a decrease of aromatic moieties from the fresh to the aged aerosol, largely due to the disappearance of phenolic structures. The lower number of carbon atoms observed in aged samples suggests that the secondary formation of oligomers was somewhat counterbalanced by fragmentation reactions and/or by the uptake of low-molecular weight compounds from the gas-phase.
- Overall, the fog composition was generally more oxidized and “SOA-like” than the aerosol, where the fresh fog composition was similar to the aged aerosol composition in both the ^1H -NMR analysis and the molecular formula trends.
- CHOS and CHNOS formulas were detected with high frequencies in samples with high water content during collection (all samples except BO0204N). This supports an enhanced production of S-containing SOA species via reactions in the aqueous phase.
- When comparing the unique formulas of the two aged samples (SPC0201F and BO0213D), aging reactions in aerosol liquid water appeared to produce less highly oxygenated CHO formulas than in fog, and a greater number of formulas in the CHNO, CHOS and CHNOS groups. This difference could be explained by the increased chance of reactions with inorganic nitrate and sulfate ions in the relatively higher solute concentrations of aerosol liquid water compared to the increased likelihood of hydration reactions in fog (Darer et al. 2011, Hu et al. 2011). This conclusion agrees with the quantitative analysis of functional group composition of aqueous SOA isolated by PMF analysis reported previously (Gilardoni et al. 2016).
- The variability of ^1H -NMR fingerprints between samples reflects the change in oxidation state of the CHO family detected by FT-ICR MS (reaching a maximum for SPC0201F), but seems rather insensitive to the changes in content of heteroatom-containing groups (CHNO, CHOS, CHNOS). In fact, the formation of CHOS compounds detected in the FT-ICR MS analysis in deliquesced aerosols (BO0213D) or in low-LWC fog water (SPC0106F) could not be traced to parallel changes in ^1H -NMR spectral characteristics. It is possible, however, that a fraction of the ^1H -NMR-detected alkoxy groups (H-C-O) were bound to sulfate esters and misclassified as alcohols.
- Compared to fresh fog (SPC0106F), the aged fog (SPC0201F) had an enhancement in the highly oxidized CHO formulas and an overall lower percentage of CHNO and CHOS formulas. This is likely due to hydrolysis reactions in the low pH environment (Darer et al. 2011, Hu et al. 2011). The ^1H -NMR analysis also highlighted that SPC0201F included highly oxidized low-molecular weight organic acids (phthalic,

maleic, succinic, pyruvic acids) which originated from the degradation of particulate WSOC, the oxidation of condensable water-soluble volatile organic compounds, and the uptake of condensable products of gas-phase oxidative reactions.

In this work, we used the detailed molecular composition to describe the differences in aging and aqueous phase processes for a select set of samples from the Supersito 2013 winter campaign. A majority of the molecular formulas observed in this study have not been previously reported, but correlate with anticipated molecular trends. This emphasizes the importance of detailed molecular analysis of atmospheric samples, for the study of biomass burning emissions processed in the aqueous phase of aerosol and fog, as well as the potential of aqueous phase processing to act as a source of SOA in the atmosphere.

2.6 Associated content

2.6.1 Data availability

An abbreviated list of the complete FT-ICR MS dataset is provided and is available on Digital Commons: <http://digitalcommons.mtu.edu/chemistry-fp/98/>

2.6.2 Author contributions

This study was designed by S. Gilardoni, S. Decesari, M. Paglione, M.C. Facchini, and L. Mazzoleni. HR-ToF-MS, ¹HNMR, and PMF analyses were done by S. Gilardoni, S. Decesari, M. Paglione. FT-ICR MS analysis was done by M. Brege and L. Mazzoleni. All authors wrote the manuscript.

2.6.3 Acknowledgments

This research was supported with a NASA Earth and Space Science Fellowship for Matthew Brege. The Italian CNR contribution was supported by the BACCHUS project, European Commission FP7-603445. The authors thank Drs. Melissa Soule and Elizabeth Kujawinski of the Woods Hole Oceanographic Institution (WHOI) Mass Spectrometry Facility for instrument time and assistance with data FT-ICR MS acquisition (NSF OCE-0619608 and Gordon and Betty Moore Foundation).

2.7 Supplementary information: chapter 2

2.7.1 FT-ICR MS data processing and molecular formula assignment review

The individual transient scans of FT-ICR MS data for each sample were reviewed manually and the unacceptable scans with an abrupt change in the total ion current were removed; the remaining transient scans were co-added together to create the working file for each

sample (this helped to increase signal to noise and enhance sensitivity). Molecular formula assignments were made as previously described (Mazzoleni et al. 2010, Putman et al. 2012, Zhao et al. 2013, Dzepina et al. 2015) using Sierra Analytics Composer software (v. 1.0.5) within the limits of: $C_{2-200}H_{4-1000}O_{1-20}N_{0-3}S_{0-1}$. Masses were calculated from measured m/z values, assuming an ion charge of -1 from the electrospray. The calculator uses a CH_2 Kendrick mass defect (KMD) analysis to sort homologous ion series and extend the molecular formula assignments to higher masses (Hughey et al. 2001, Kujawinski and Behn 2006). A de novo cut-off at m/z 500 was applied, indicating that no new formula assignments would occur above m/z 500, unless the formula was part of an existing CH_2 homologous series that began at a point lower than m/z 500. This is necessary because the number of possible molecular formulas increases at higher values. The minimum relative abundance required for molecular formula assignment was > 10 times the estimated signal-to-noise ratio, determined for each sample between m/z 900–1000. Only integer values up to 40 were allowed for the double bond equivalents (DBE). The data set was manually reviewed to remove: formulas with an absolute error > 3 ppm, elemental ratios that were not chemically sensible (such as $O:C > 3$ or $H:C < 0.3$), and formulas which violated the rule of 13 or violated the nitrogen rule. The rule of 13 checks for a reasonable number of heteroatoms in a formula. A base formula (C_nH_{n+r}) can be generated for any measured mass by solving: $\frac{M}{13} = n + \frac{r}{13}$ (Pavia 2009). Then, the maximum number of "large atoms" (C, O, N, S) in a formula is defined as the mass divided by 13, because substituting for a heteroatom (O, N or S) involves a substitution for at least one carbon. This maximum number is then compared to the actual number of "large atoms" in a formula, and those formulas exceeding the maximum number are rejected. The nitrogen rule removes formulas with odd masses that do not contain an odd number of nitrogen atoms, and even masses that do not contain an even number (or no) nitrogen atoms; this is due to the odd numbered valence of nitrogen (Pavia 2009). Molecular formulas that contained ^{13}C or ^{34}S were also removed from the data set. Homologous series with large gaps in the DBE trend were removed, as well as homologous series with a length of one. The assigned formulas were also analyzed with consideration to the DBE and oxygen number trends, (Herzprung et al. 2014) where unreliable formula assignments were also removed.

2.7.2 Ultrahigh resolution FT-ICR MS results

The total ion abundance of the identified monoisotopic molecular formulas reported for each sample was determined by their summation. Then, these values were used to normalize the individual ion abundances within each sample using a ratio of the individual ion intensity to this total ion abundance. Then, the values were rescaled using a normalization constant (10,000). This normalization procedure was done to remove analytical biases introduced by trace contaminants with high electrospray efficiency.

Reconstructed difference mass spectra of the assigned molecular formulas for both fog and aerosol samples are shown in Fig. 2.8. These difference mass spectra permit a direct comparison of the samples using normalized relative abundances. The individual relative abundances were normalized by the total abundance of the assigned molecular formulas

identified in each of the samples. In Fig. 2.8, the individual masses with higher abundances in either the positive or negative direction were substantially greater in one of the two samples, whereas the masses of similar abundance tended to cancel each other. To enhance the interpretation of the compositional differences, the individual masses were color-coded to represent the number of oxygen atoms in the assigned formula. Overall, we observed higher numbers of oxygen in the masses of the two samples with aged biomass burning emissions influence compared to the two samples with fresh biomass burning emissions influence. The molecular formulas assigned to the fresh samples had approximately 0 to 5 oxygen atoms over the mass range of 50 to 250 Da, 5 to 10 oxygen atoms over 250 to 550 Da, and a few molecular formulas were assigned with 10 to 15 oxygen atoms over 500 to 600 Da. In contrast, the aged samples had a large number of molecular formulas with 10 to 15 oxygen atoms in the range of 400 to 550 Da. This clearly shows a greater amount of oxidation in the aged influenced samples compared to the fresh influenced samples.

KMD diagrams can be used as useful tools to visualize the relationships between the many molecular formulas of complex mixtures such as atmospheric samples. We used Kendrick mass defect to sort the molecular formulas into CH_2 homologous series of identical heteroatom content and DBE, where the formulas in the same series differ only by a number of CH_2 units (Stenson et al. 2003). It should be noted that the presence of multiple formulas in the same homologous series does not necessarily imply a related chemical structure. The homologous series are visible as horizontal rows of formulas in Fig. 2.12. There were multiple homologous series per subclass, where the base formula for each series differ in DBE and increase in KMD to form an ensemble of “steps” within each subclass. In our samples individual CHO and CHNO subclasses had approximately 5 to 16 different homologous series, while CHOS and CHNOS subclasses had approximately 3 to 10 different homologous series. The number of homologous series in a subclass increased with oxygen number, and peaked near the median oxygen number, then decreased again towards the maximum number of oxygen; this led to fewer molecular formulas in subclasses with higher and lower oxygen numbers, and more formulas in subclasses near the median oxygen number. The subclasses with the highest numbers of molecular formulas per elemental group were: O_7 , NO_8 , O_7S and NO_9S . It was atypical for the unique formulas of a sample to be completely unrelated to other formulas across the data set; often the unique formulas were extensions of homologous series that appeared across samples.

3 Molecular trends observed in sequential samples of biomass burning-influenced aerosol and fog water in the Po Valley analyzed by ultrahigh resolution Orbitrap Elite mass spectrometry²

3.1 Abstract

Biomass burning emissions and aqueous phase processing of atmospheric aerosol each contribute substantially to the overall complexity observed in the organic fraction. The detailed molecular characterization of aerosol impacted by both of these processes is somewhat limited in the scientific literature. In this study, we collected concurrent aerosol and fog water samples in the Po Valley over a four-day period in the winter of 2015. The calculated absorption Ångström exponent for the samples suggest a significant wavelength dependence on light absorption and the presence of brown carbon. Our comprehensive analysis by ultrahigh resolution Orbitrap Elite mass spectrometry using positive and negative mode electrospray ionization revealed extensive molecular diversity. We observed a significant influence of molecular species with nitrogen (~53%). We expect oxidized organonitrates from CHNO group formulas to be present in most samples and reduced imine-like and imidazole-like species from CHN group formulas, which were observed in greater intensity in night time samples. Organosulfate and nitro-oxy organosulfate species were also observed in daytime aerosol samples with enhanced intensity. Fog water samples trended towards lower carbon numbers compared to aerosol, likely related to the enhanced oxidation and fragmentation expected from chemical transformation in the aqueous phase. The modified aromaticity index of the molecular formulas predicted a significant number of aromatic and condensed aromatic species (~22%), and we suspect that these species which also contain nitrogen are related to brown carbon. The ambient samples were compared to a laboratory sample containing pyrazine-based chromophores prepared via reactions between methyl glyoxal and ammonium sulfate. There was significant overlap (~16 to 23%) between the molecular formulas observed in the ambient and laboratory samples. Given the high calculated AAE for the samples, the significant fraction of formulas with nitrogen and estimated aromatic character, we suspect that pyrazine-based chromophores may be present in these ambient samples. The molecular trends observed here help to describe the fundamental transformations of biomass burning emissions by aqueous phase processing in fog and aerosol liquid water. The aromatic and nitrogen containing molecular formulas observed here may also be relevant to the fraction of brown carbon expected to be produced by biomass burning and by fog processing.

²As stated in the Preface, the material contained in chapter 3 was prepared with the intent of submission for publication. As of 9-Dec 2019 the manuscript has not yet been submitted.

3.2 Introduction

The molecular diversity of chemical species observed in the organic fraction of atmospheric aerosol is the product of numerous coexisting atmospheric processes (Fuzzi et al. 2015, Poschl and Shiraiwa 2015, Shrivastava et al. 2017). Emissions from biomass combustion using woody, vegetative and animal waste fuels contributes significantly to the organic fraction of atmospheric aerosol (Andreae and Merlet 2001, Bond et al. 2004, Glasius et al. 2006, Fleming et al. 2018). Chemical reactions in the aqueous phase can lead to secondary transformation products which are unexplained by gas phase oxidants and photolysis (Kanakidou et al. 2005, Ervens et al. 2011, McNeill 2015) and the transformation of biomass burning emissions has been observed to have important implications for air quality and climate (Gilardoni et al. 2016). Liquid water droplets are efficient scavengers of water-soluble species including organics (Facchini et al. 1999, Kampf et al. 2013, Gilardoni et al. 2014) and co-solvation effects are expected to aid in the dissolution of marginally soluble species (Dickhut et al. 1989, Fan and Jafvert 1997, Wang et al. 2019). Some aerosol are expected to remain inactivated in the interstitial space between droplets (Gilardoni et al. 2014) further complicating the diversity in atmospheric chemical processes. In our previous work (see chapter 2) we compared the aqueous phase processes occurring in aerosol liquid water to fog droplets (Brege et al. 2018). In the assignment of molecular formulas to aerosol and fog samples with diverse influences from “fresh” emissions and “aged” processes, as well as liquid water content, we observed several molecular trends. Molecular formulas in fog had lower carbon numbers and higher oxygen numbers compared to those observed in aerosol. The abundance of organosulfate species was observed to increase in aerosol samples with high liquid water content compared to dry aerosol and in low liquid water content fog compared to fog with excess liquid water. The organosulfates were expected to have formed via chemical reactions in the relatively high concentrations of aerosol liquid water and then degraded via hydrolysis reactions in fog droplets (Darer et al. 2011, Hu et al. 2011). Multiple cycles of aqueous phase processing as droplets dry and reform during the lifetime of a fog can intensify these chemical transformations on the organics in aerosol (Lee et al. 2012).

Light absorbing chemical species in atmospheric aerosol have a significant impact on the overall global climate, as they can lead to an overall increased absorption of solar radiation and cause atmospheric heating (Kanakidou et al. 2005, Moosmuller et al. 2009, IPCC 2018). Light absorbing carbonaceous species are generally described as “black” or “brown” carbon, where black carbon has a refractory graphitic chemical structure which nearly universally absorbs solar radiation independent of wavelength, and brown carbon is a poorly characterized chemically mixture of smaller organic species that more selectively absorb solar radiation at higher wavelengths in the blue and ultra-violet regions (Andreae and Gelencser 2006, Moosmuller et al. 2009). This wavelength dependence on light absorption for aerosol can be calculated via the absorption Ångström exponent (AAE), where light absorption due entirely to black carbon will result in an $AAE \approx 1$ and light absorption from brown carbon will result in calculated AAE values much greater than 1 (Moosmuller et al. 2009). The sources and lifetimes of brown carbon species in the atmosphere are still somewhat uncertain, and organic species directly emitted (Chakrabarty

et al. 2010, Desyaterik et al. 2013, Saleh et al. 2018) as well as those expected to form as secondary species (Lee et al. 2013, Nguyen et al. 2013, Hawkins et al. 2018, Huang et al. 2018) both contribute to the total absorption by brown carbon. Organonitrates and organic amines may play an important role in the chemical composition of brown carbon (Desyaterik et al. 2013, Lee et al. 2013, Hawkins et al. 2018, Adachi et al. 2019), as well as functionalized methoxy-phenols (Yu et al. 2014, Yu et al. 2016) and polycyclic organic hydrocarbons (Lin et al. 2018, Saleh et al. 2018). Field observations (Gilardoni et al. 2016) and laboratory experiments (Lee et al. 2013, Hawkins et al. 2018, Huang et al. 2018) alike suggest that aqueous phase processes can lead to enhanced light absorption by the formation of brown carbon species. Costabile et al. (2017) describe the correlation of brown carbon species to high concentrations of NH_4^+ and NO_3^- in aerosol as well as low ratios of black carbon to the total organic aerosol mass.

The Po Valley in the Emilia-Romagna region of Italy is an area strongly impacted by biomass burning pollution from residential wood combustion and agriculture (Pietrogrande et al. 2015). Frequent fog formation occurs in the valley during the winter season, where aqueous phase processes are expected to influence the regional aerosol. The concentration of organic species in Po Valley fog water are among the highest reported in the world (Herkes et al. 2013). Thus, the chemical analysis of aerosol from this region has the potential to reveal compelling molecular level details on the nature of brown carbon containing ambient aerosol. Numerous field campaigns have been conducted since 1989 (Giulianelli et al. 2014) and the regional influences on aerosol composition have been well studied. Real time chemical characterization of aerosol using a high resolution-time of flight-aerosol mass spectrometer has been a part of these field campaigns (Gilardoni et al. 2014, Dall'Osto et al. 2015) and further detailed molecular characterization of Po Valley fog and aerosol using ultrahigh resolution-Fourier transform ion cyclotron resonance mass spectrometry has also been described in our previous work (Brege et al. 2018). Thus far, because of the significant time investment involved in ultrahigh resolution mass spectrometry data analysis, and the extreme complexity of the organic matter observed, the detailed molecular characterization of aerosol for the Po Valley has been conducted on a limited number of samples and analysis methods.

In this study we have analyzed a set of 11 concurrent fog water and aerosol samples collected over a four-day period of the 2015 Po Valley winter field campaign by ultrahigh resolution Orbitrap Elite mass spectrometry. We have performed the analyses with both the positive and negative polarities of “soft-ionization” electrospray ionization, to increase the observable detail and reveal the most molecular formulas possible. We have developed semi-automated scripts for the post-processing and molecular formula assignment of the resulting large data set. Here we related the detailed molecular compositions of these samples to one another, to ancillary measurements collected during the field campaign and a laboratory generated aqueous secondary organic aerosol sample. Here we describe molecular species which we conclude to be important contributors to atmospheric brown carbon and that have not been observed previously.

Figure 3.1. Measured (blue) and modeled (red) liquid water content (LWC) for fog droplets and aerosol respectively over the sample collection period between 1-Dec and 4-Dec 2015. Sample collection times are indicated with colored rectangles. Note the difference in scale between LWC measured for fog (g m^{-3}) and modeled for aerosol ($\mu\text{g m}^{-3}$). Aerosol LWC was estimated using the ISORROPIA-II model based on field campaign data.

3.3 Methods

Sub-micron aerosol (PM_{10}) were collected from approximately 9:00 to 18:00 for “day” samples and from 18:00 to 9:00 (the following morning) for “night” samples. Aerosol samples were collected on pre-baked quartz fiber filters using a high-volume sampler (TECORA Echo Hi Vol) equipped with a digital PM_{10} sampling inlet, at a nominal flow rate of $30 \text{ m}^3 \text{ h}^{-1}$. Fog sampling occurred at night by an automatic fog collection system (Fuzzi et al. 1997) where droplets are pulled through a short wind tunnel and coalesce by impaction onto stainless steel strings then flow into a collection bottle. Samples were collected from 1-Dec 2015 to 4-Dec 2015 at San Pietro Capofiume (SPC) (Emilia-Romagna, Italy) during the 2015 winter field campaign. As a part of the campaign several continuous online ambient measurements were performed including droplet liquid water content (LWC) (Fig. 3.1), and chemical composition via high resolution-time of flight-aerosol mass spectrometry (HR-ToF-AMS) (Aerodyne Research (Gilardoni et al. 2014)). Likewise, aliquots of fog samples were analyzed by ion chromatography to determine the concentration of common inorganic ions (Cl^- , NO_3^- , SO_4^{2-} , Na^+ , NH_4^+ , K^+ , Mg^{2+} and Ca^{2+}) and low molecular weight organics (acetate, formate, methanesulfonic acid, oxalate, methylamine, ethylamine, dimethylamine, diethylamine and trimethylamine). Offline absorbance measurements of aliquots of the fog samples and water extracted aerosol filter portions were measured by a flow cell UV-vis spectrophotometer (TIDAS) with a 50 cm path length. Aerosol LWC was calculated by the ISORROPIA-II model based on field campaign data (Fig. 3.1).

Figure 3.2. An example of mass spectra stitching for the SPC03 daytime (-ESI) sample. The paired spectra collected at different scan ranges (a & b) were compared to one another post formula assignment to determine common molecular formulas from m/z 200 to 400 (cyan lines). The average ratio of these common formulas between the two data sets (panel a:panel b) was used to combine the spectra (c) and normalize the intensities for the final version (d).

A portion of the aerosol filters and an aliquot of the fog samples were transported frozen to Houghton, MI (USA) and stored frozen until ultrahigh resolution Orbitrap Elite mass spectrometry (OTE-MS) sample analysis could be performed. The collected samples are referred to here as “SPC01”, “SPC02”, “SPC03” and “SPC04” with respect to their sampling dates and location; daytime, nighttime and fog samples will be referred to specifically in context. The SPC04 nighttime aerosol filter sample was exposed to excess water in the sampler and thus does not represent a sample consisting exclusively of aerosol, so it was removed from the sample set. Thus, in total there were 11 samples for analysis from this four-day period. An aerosol filter blank and a sample preparation blank were also prepared in tandem with the other samples for OTE-MS analysis. In preparation for analysis by OTE-MS, the aerosol filter portions were cut into small pieces and extracted in 15 mL of high purity water using an orbital shake table for 90 min at 60 RPM. The extracted aerosol samples and the fog sample aliquots were then filtered by a 0.2 μm PTFE syringe filter to remove insoluble material. The extracted samples underwent a solid phase extraction procedure to remove soluble salts, concentrate the analytes and exchange to an organic solvent more suitable for OTE-MS analysis. This SPE procedure used reverse-phase Oasis HLB (6 cc 200 mg) cartridges. The aerosol filter extracts were eluted off the HLB cartridges with 3 mL of 10% water in acetonitrile solution and the fog extracts were eluted with 3 mL of 10% water in methanol solution. The molecular characterization presented here is thus influenced by this aqueous phase extraction. Though there is the potential for co-solvation effects from lower-polarity species present in the mixture (Dickhut et al. 1989, Fan and Jafvert 1997) a majority of the chemical species observed here are expected to be water-soluble organic compounds.

The sample extracts were analyzed in the Chemical Advanced Resolution Methods (ChARM) Lab at Michigan Technological University using OTE-MS. The samples were directly infused at 5 to 10 $\mu\text{L min}^{-1}$ and analyzed using the positive and negative ion modes of ESI. These ionization modes will be referred to as +ESI and -ESI from this point forward. Spectra were collected for 100 to 200 scans at a resolution of 240,000 (defined at m/z 400) with spectrum averaging on. The extracts were analyzed from m/z 100 to 800 and m/z 200 to 800, in order to maximize the ion signal for species which may be suppressed by efficiently ionizable analytes at lower m/z (Cech and Enke 2001). For each sample the average ratio of intensity observed in the m/z 100 to 800 to the m/z 200 to 800 scan range was determined for formulas appearing in both scan ranges over m/z 200 to 400 (see Fig. 3.2). This ratio was used to normalize the intensity of both scan ranges and combine the spectra. As the measured intensity of ions in ESI is non-quantitative due to charge competition during ionization, this combination of scan ranges provides significant additional molecular detail at higher masses with only a minor decrease in confidence regarding ion intensities. This spectral stitching was performed for reconstructed mass spectra after formula assignment to avoid the need to average measured masses.

Mass lists were generated from the collected spectra using Xcalibur software (Thermo Scientific v. 3.0). Molecular formulas were assigned to the mass lists using the R package MFAssignR (v. 0.0.4) (Schum et al. 2018), using a series of custom scripts developed for R and R Studio (v. 3.5.1 and v. 1.2.1335). A minimum intensity of 6 times the estimated

signal to noise was required for formula assignment. C, H, N, O and S elements were allowed, with limits of N₀₋₃, S₀₋₁ and the default limits of O₀₋₃₀. Up to one Na atom was allowed for assignments in +ESI to account for expected Na⁺ adducts in addition to H⁺ adducts. These restrictions on heteroatoms severely reduces the number of incorrect formula assignments in the resulting data set. A preliminary formula assignment specific to each sample (allowing only CHO species) was used to determine high intensity CH₂ homologous series to use as seed series for the recalibration function built in to MFAssignR. This recalibration of the spectra reduced the overall error in formula assignment and allowed for additional formulas to be assigned. Molecular formulas were blank subtracted based on the top 10% of ions detected in the sample preparation blanks. This was done to remove obvious background signal ions from the samples and to avoid the gratuitous removal of molecular formulas which may be falsely detected in the blank sample analyses due to the low ion signal associated with blank sample analyses. Samples analyzed by +ESI were compared to the full blank spectrum above *m/z* 400 due to high instrument background signal. The assigned molecular formulas were passed through a robust validation process to ensure data quality, similar to our previous works (Brege et al. 2018, Schum et al. 2018). Intensities were normalized to the total ion intensity of the spectra for each sample in order to reduce the influence of a small number of high intensity peaks on the entire spectra. Unless stated otherwise the ion intensities discussed in this work were normalized in this way.

The double bond equivalents (DBE) (unsaturated bonds and/or ring structures) for a molecular formula can be calculated based on the number of missing electrons expected, though DBE does not account for which atoms the unsaturation may be between. The modified aromaticity index (AI_{mod}) was used to estimate the carbon-carbon bond saturation and predict aromatic character of the molecular formulas Koch and Dittmar (2006, 2016). The calculation for DBE is provided in equation (3.1) and the calculation for AI_{mod} is described in equations (3.2) to (3.4). Equations (3.1) to (3.4) use the molecular formula format: C_cH_hN_nO_oS_s.

$$DBE = c - \left(\frac{h}{2}\right) + \left(\frac{n}{2}\right) + 1 \quad (3.1)$$

$$DBE_{AI,mod} = 1 + c - \frac{o}{2} - s - \left(\frac{n+h}{2}\right) \quad (3.2)$$

$$C_{AI,mod} = c - \frac{o}{2} - n - s \quad (3.3)$$

$$AI_{mod} = \frac{DBE_{AI,mod}}{C_{AI,mod}} \quad (3.4)$$

In equation (3.4), the AI_{mod} = 0, if DBE_{AI,mod} ≤ 0 or C_{AI,mod} ≤ 0, as defined by Koch and Dittmar (2006, 2016). In the AI_{mod} calculation for N and S-containing formulas from -ESI analyses, equations (3.2) and (3.3) were adjusted to account for a number of oxygen atoms in the formula which are assumed to be bound in nitro, nitrate, sulfonate and sulfate groups, by substituting the $\left(\frac{o}{2}\right)$ term for $\left(\frac{o}{2} - 3\right)$. AI_{mod} values are defined by Koch and Dittmar

(2006, 2016) as follows: aliphatic ($AI_{\text{mod}} = 0$), olefinic ($0 < AI_{\text{mod}} \leq 0.5$), aromatic ($AI_{\text{mod}} > 0.5$) and condensed aromatic ($AI_{\text{mod}} \geq 0.67$). In this work we also required condensed aromatics to have at least 8 carbon atoms to avoid misclassifying formulas which could not reasonably have multiple ring structures.

3.4 Results and discussion

The results of the ambient measurements from the Po Valley winter 2015 field campaign are summarized in Table 3.1, including measured and modeled LWC, the calculated wavelength dependence on aerosol light absorption AAE and inorganic ion (NO_3^- , NH_4^+ and SO_4^{2-}) concentrations for the respective sample collection times. The AAE was calculated based on UV-vis data over a range of 360 to 600 nm wavelengths, and the resulting values for these samples show significant influence from wavelength dependent light absorption by brown carbon ($AAE \gg 1$). The inorganic ions present are expected to react with the organic matter in aerosol and fog to form organosulfates and organonitrates (Brege et al. 2018) as well as organic amines (Malloy et al. 2009). LWC content as a function of sample collection time is presented in Fig. 3.1, including the measured LWC for droplets as well as the modeled LWC for aerosol.

Approximately 4000 to 8000 molecular formulas were observed in each sample and in general, more formulas were observed in -ESI compared to +ESI. A summary of the molecular formula assignment for each sample are provided in Tables 3.2 and 3.3. The reconstructed mass spectra for the samples are shown in Appendix A (Figs. 6.1 to 6.11). The overall percent composition of elemental groups in each sample are shown in Fig. 3.3. These elemental groups (“CHO”, “CHNO”, “CHOS”, “CHNOS”, “CHN” and “CH”) describe the basic elemental composition of the molecular formula. There were a significant number of nitrogen species in the form of CHNO (~50% of formulas) and a minor amount of CHN species. There were moderate amounts of sulfur species (~10 to 20% of formulas) in about equal amounts of CHOS and CHNOS species, the latter of which further contributes to the enhanced influence of nitrogen in these samples.

Daytime samples had CHOS formulas with particularly high ion intensities, which were further validated by matching to the expected ^{34}S isotope peaks (Fig 3.4). There was a slightly higher percent of CHO formulas observed in night samples compared to the day and fog samples. There was increased detection of CHOS and CHNOS formulas in the -ESI analyses and increased detection of CHNO in the +ESI analyses. We expect that reduced and less polar species are more efficiently ionized in +ESI compared to -ESI (Cech and Enke 2001), and we observe a small number of CH and CHN formulas in these analyses (< 5% of formulas). Overall, these results are generally consistent with what was observed in our previous work for the molecular characterization of Po Valley fog and aerosol. The increased influence from CHNO species is observed here as a result of the inclusion of analyses by +ESI in addition to -ESI.

Table 3.1. Summary of Po Valley winter 2015 field campaign ambient measurements.

		AAE ^a	LWC ^b	[SO ₄ ²⁻] ^c	[NH ₄ ⁺] ^c	[NO ₃ ⁻] ^c
SPC01	Day	6.26	60.1 ug m ⁻³	2.09 ug m ⁻³	4.55 ug m ⁻³	11.9 ug m ⁻³
	Night	6.31	34.0 ug m ⁻³	0.917 ug m ⁻³	2.02 ug m ⁻³	5.48 ug m ⁻³
	Fog	4.54	0.0775 g m ⁻³	8.14 mg L ⁻¹	24.23 mg L ⁻¹	66.11 mg L ⁻¹
SPC02	Day	6.39	64.4 ug m ⁻³	1.47 ug m ⁻³	4.08 ug m ⁻³	11.3 ug m ⁻³
	Night	6.14	32.0 ug m ⁻³	0.550 ug m ⁻³	1.00 ug m ⁻³	2.51 ug m ⁻³
	Fog	4.78	0.138 g m ⁻³	4.82 mg L ⁻¹	12.49 mg L ⁻¹	29.47 mg L ⁻¹
SPC03	Day	6.33	49.1 ug m ⁻³	0.837 ug m ⁻³	1.52 ug m ⁻³	3.92 ug m ⁻³
	Night	5.73	14.3 ug m ⁻³	0.369 ug m ⁻³	0.403 ug m ⁻³	0.812 ug m ⁻³
	Fog	5.18	0.124 g m ⁻³	5.22 mg L ⁻¹	10.82 mg L ⁻¹	24.22 mg L ⁻¹
SPC04	Day	6.10	16.4 ug m ⁻³	0.415 ug m ⁻³	0.486 ug m ⁻³	1.05 ug m ⁻³
	Night	6.90	9.46 ug m ⁻³	0.288 ug m ⁻³	0.271 ug m ⁻³	0.585 ug m ⁻³
	Fog	4.57	0.0597 g m ⁻³	10.73 mg L ⁻¹	24.33 mg L ⁻¹	74.66 mg L ⁻¹

^aAbsorption Ångström exponent (AAE) was determined for offline aerosol filter and fog samples via UV-vis spectroscopy, over wavelengths 360 to 600 nm. ^bLiquid water content (LWC) for fog samples was measured as a part of the field campaign and estimated for aerosol by the ISORROPIA-II model (Note the difference in units between sample types).

^cInorganic salt concentrations in aerosol were determined via online high resolution-time of flight-aerosol mass spectrometry and were determined for fog via offline ion chromatography (Note the difference in units between sample types).

Elemental ratios based on the molecular formulas are shown in Fig. 3.9. As expected, O/C was shifted higher in -ESI and H/C was shifted higher in +ESI. The estimation of carbon-carbon bond saturation provided by AI_{mod} showed ~10 to 15% of formulas were estimated to be aliphatic and ~20 to 30% of formulas were estimated to be aromatic or condensed aromatic. There was an enhanced percentage of formulas which were estimated to be aliphatic in +ESI analyses, as well as a decreased percentage of estimated aromatics and condensed aromatics. There was an overall increasing trend for the percentage of estimated aliphatic species from 1-Dec to 4-Dec in the +ESI analyses.

Carbon intensity trends based on the normalized intensities of the assigned molecular formulas are provided in panels a and b of Figs. 3.6 to 3.15. These trends represent the total measured ion intensity of molecular formulas binned for a carbon number (C_n). The difference of these intensity trends are shown in panels c and d of the same Figures. Some repetition is present in Figs. 3.6 to 3.15 from the comparisons between day-night, day-fog and fog-night samples (partially excluding SPC04). These results are overall consistent with our previous results (Brege et al. 2018) for Po Valley fog and aerosol. Most samples had increased intensities between C_5 and C_{15} and log-normal-like decreasing trends towards higher carbon numbers. In many samples there was an enhanced intensity for C_{10} which may be influence from biogenic monoterpenes, which have been observed in biomass burning emissions previously (Zhao et al. 2013, Cook et al. 2017, Brege et al. 2018). There was an enhanced intensity for CHOS and CHNOS species in the -ESI analyses, especially for daytime samples, which could be related to secondary reactions between organics and sulfate salts in aerosol liquid water (Noziere et al. 2010, Darer et al. 2011, Hu et al. 2011, McNeill et al. 2012). There was an increased amount of aerosol liquid water during the daytime sample collection especially near solar noon (Fig. 3.2).

The high concentrations of salts in aerosol liquid water are expected to lead to different transformation products than in the comparably dilute concentrations of fog, where more liquid water is present (Darer et al. 2011, Hu et al. 2011). These organosulfate species are susceptible to hydrolysis reactions (Darer et al. 2011, Hu et al. 2011) and would thus be expected to degrade in fog droplets, as observed here by the comparably lower intensity of CHOS and CHNOS species in fog compared to daytime samples. CHN species were observed with greater intensity in nighttime samples compared to daytime aerosol and fog, though there is a unique “even” carbon number preference for CHN species in fog samples observed at C_6 , C_8 and C_{10} . This pattern is somewhat present in daytime samples as well, although marginally diminished, possibly due to influence from the preceding nighttime aerosol. As stated previously there was an enhanced observation of CHNO molecular formulas in +ESI analyses which can also be observed in these carbon intensity trends.

The difference of these intensity trends (panels c and d of Figs. 3.6 to 3.15) is discussed in the following subsections, including day-night, day-fog and fog-night trends. In these difference intensity trends the differences between samples become more apparent, as molecular formulas appearing in both samples with similar intensities are reduced to zero and species which are more abundant in one sample or another have enhanced intensities.

Table 3.2. Summary of molecular formula assignment for +ESI analyses of SPC samples from the 2015 winter field campaign.

		All	CHO	CHNO	CHOS	CHNOS	CH	CHN	
SPC01	Day	Na	6570	1274	5096	60	0	22	118
		O/C ^b	0.121	0.169	0.106	0.633	0	0	0
		H/C ^b	1.254	1.205	1.256	1.967	0	1.273	1.331
		DBE ^{b,c}	9.022	8.092	9.409	2.833	0	5.727	6.127
		AI _{mod} ^{b,c}	0.179	0.281	0.150	0	0	0.182	0.407
SPC01	Fog	Na	5365	1176	3984	27	62	21	95
		O/C ^b	0.125	0.162	0.109	0.741	0.403	0	0
		H/C ^b	1.281	1.27	1.270	1.889	1.935	1.238	1.253
		DBE ^{b,c}	8.241	7.577	8.610	3.889	2.984	5.476	6.274
		AI _{mod} ^{b,c}	0.200	0.289	0.173	0.037	0	0.143	0.421
SPC01	Night	Na	5600	2046	3365	46	4	20	119
		O/C ^b	0.117	0.147	0.095	0.717	0	0	0
		H/C ^b	1.261	1.207	1.286	1.761	1.000	1.250	1.294
		DBE ^{b,c}	8.834	9.534	8.585	3.522	13.000	5.400	6.319
		AI _{mod} ^{b,c}	0.184	0.192	0.173	0	1.000	0.150	0.412
SPC02	Day	Na	6414	1875	3894	370	135	24	116
		O/C ^b	0.158	0.173	0.112	0.519	0.459	0	0
		H/C ^b	1.361	1.291	1.346	1.822	1.541	1.208	1.353
		DBE ^{b,c}	7.731	8.181	8.086	3.730	4.170	5.750	5.871
		AI _{mod} ^{b,c}	0.145	0.175	0.141	0	0.030	0.208	0.353
SPC02	Fog	Na	4466	1050	3176	25	114	19	82
		O/C ^b	0.128	0.180	0.104	0.600	0.342	0	0
		H/C ^b	1.322	1.330	1.303	1.840	1.702	1.263	1.244
		DBE ^{b,c}	7.542	6.991	7.946	2.280	4.070	5.105	5.939
		AI _{mod} ^{b,c}	0.184	0.259	0.161	0	0.035	0.158	0.390
SPC02	Night	Na	5244	1520	3309	155	118	16	126
		O/C ^b	0.167	0.207	0.117	0.703	0.542	0	0
		H/C ^b	1.334	1.297	1.316	1.884	1.636	1.188	1.317
		DBE ^{b,c}	7.845	8.033	8.192	3.252	3.712	5.063	6.325
		AI _{mod} ^{b,c}	0.183	0.210	0.178	0	0	0.250	0.389

^aValues represent the respective number of molecular formulas. ^bValues represent mathematical averages based on formula assignment. ^cDouble bond equivalents (DBE) and modified aromaticity index (AI_{mod}) values were obtained using equations (3.1) and (3.2) to (3.4) respectively.

(Table 3.2 continued next page)

(Table 3.2 continued from previous page)

		All	CHO	CHNO	CHOS	CHNOS	CH	CHN	
SPC03	Day	Na	6300	2022	3708	310	121	23	116
		O/C ^b	0.174	0.175	0.140	0.539	0.471	0	0
		H/C ^b	1.337	1.257	1.327	1.858	1.686	1.174	1.336
		DBE ^{b,c}	8.027	8.791	8.187	3.587	3.711	5.913	6.414
		AI _{mod} ^{b,c}	0.146	0.176	0.137	0	0.025	0.217	0.397
SPC03	Fog	Na	3938	989	2231	357	270	17	74
		O/C ^b	0.181	0.180	0.082	0.622	0.485	0	0
		H/C ^b	1.429	1.376	1.358	1.838	1.763	1.059	1.162
		DBE ^{b,c}	6.356	6.664	7.147	2.779	3.504	5.882	6.162
		AI _{mod} ^{b,c}	0.171	0.281	0.160	0	0.007	0.353	0.432
SPC03	Night	Na	5013	1529	2458	590	301	28	107
		O/C ^b	0.266	0.224	0.181	0.592	0.654	0	0
		H/C ^b	1.444	1.335	1.395	1.802	1.797	1.107	1.262
		DBE ^{b,c}	6.606	7.417	7.245	3.434	3.538	6.464	6.477
		AI _{mod} ^{b,c}	0.161	0.205	0.177	0	0.003	0.500	0.430
SPC04	Day	Na	5247	1645	3013	335	122	26	106
		O/C ^b	0.186	0.212	0.128	0.588	0.377	0	0
		H/C ^b	1.355	1.292	1.333	1.830	1.525	1.115	1.321
		DBE ^{b,c}	7.42	7.887	7.805	3.364	4.189	6.038	6.142
		AI _{mod} ^{b,c}	0.148	0.167	0.149	0	0.049	0.423	0.358
SPC04	Fog	Na	5892	1305	4382	51	39	21	94
		O/C ^b	0.160	0.229	0.135	0.608	0.462	0	0
		H/C ^b	1.297	1.267	1.297	1.745	1.615	1.333	1.330
		DBE ^{b,c}	8.112	7.514	8.424	4.451	5.103	5.333	5.670
		AI _{mod} ^{b,c}	0.148	0.225	0.121	0.157	0.231	0.143	0.330

^aValues represent the respective number of molecular formulas. ^bValues represent mathematical averages based on formula assignment. ^cDouble bond equivalents (DBE) and modified aromaticity index (AI_{mod}) values were obtained using equations (3.1) and (3.2) to (3.4) respectively.

Table 3.3. Summary of molecular formula assignment for -ESI analyses of SPC samples from the 2015 winter field campaign.

		All	CHO	CHNO	CHOS	CHNOS	
SPC01	Day	N ^a	8303	3029	3659	939	676
		O/C ^b	0.355	0.259	0.303	0.596	0.725
		H/C ^b	1.209	1.155	1.109	1.544	1.528
		DBE ^{b,c}	9.49	10.516	10.71	4.677	4.969
		AI _{mod} ^{b,c}	0.249	0.203	0.359	0.084	0.087
SPC01	Fog	N ^a	7129	2290	3052	1075	712
		O/C ^b	0.293	0.202	0.21	0.502	0.626
		H/C ^b	1.245	1.207	1.119	1.526	1.487
		DBE ^{b,c}	8.685	9.581	9.926	5.27	5.646
		AI _{mod} ^{b,c}	0.322	0.296	0.471	0.095	0.111
SPC01	Night	N ^a	7596	2819	3649	691	437
		O/C ^b	0.332	0.242	0.301	0.588	0.769
		H/C ^b	1.215	1.168	1.13	1.579	1.638
		DBE ^{b,c}	9.153	10.202	9.895	4.184	4.048
		AI _{mod} ^{b,c}	0.265	0.196	0.376	0.094	0.057
SPC02	Day	N ^a	8347	3167	3244	1257	679
		O/C ^b	0.248	0.157	0.201	0.39	0.63
		H/C ^b	1.269	1.195	1.153	1.533	1.68
		DBE ^{b,c}	9.127	10.505	10.282	5.321	4.224
		AI _{mod} ^{b,c}	0.232	0.193	0.364	0.083	0.06
SPC02	Fog	N ^a	6714	2076	2555	1177	906
		O/C ^b	0.299	0.195	0.178	0.522	0.588
		H/C ^b	1.252	1.186	1.102	1.531	1.461
		DBE ^{b,c}	8.909	10.323	10.445	5.218	6.129
		AI _{mod} ^{b,c}	0.339	0.347	0.537	0.082	0.094
SPC02	Night	N ^a	8057	2946	3417	981	713
		O/C ^b	0.304	0.203	0.241	0.529	0.714
		H/C ^b	1.233	1.158	1.146	1.518	1.571
		DBE ^{b,c}	9.296	10.824	10.168	4.926	4.818
		AI _{mod} ^{b,c}	0.247	0.204	0.351	0.127	0.095

^aValues represent the respective number of molecular formulas. ^bValues represent mathematical averages based on formula assignment. ^cDouble bond equivalents (DBE) and modified aromaticity index (AI_{mod}) values were obtained using equations (3.1) and (3.2) to (3.4) respectively.

(Table 3.3 continued next page)

(Table 3.3 continued from previous page)

		All	CHO	CHNO	CHOS	CHNOS	
SPC03	Day	N ^a	8206	3023	3271	1136	776
		O/C ^b	0.294	0.184	0.246	0.438	0.713
		H/C ^b	1.275	1.176	1.172	1.585	1.647
		DBE ^{b,c}	9.12	10.802	10.113	4.905	4.552
		AI _{mod} ^{b,c}	0.23	0.208	0.337	0.095	0.062
SPC03	Fog	N ^a	5170	1863	2170	686	451
		O/C ^b	0.246	0.174	0.181	0.446	0.552
		H/C ^b	1.316	1.254	1.212	1.641	1.576
		DBE ^{b,c}	7.824	8.835	8.744	4.085	4.905
		AI _{mod} ^{b,c}	0.326	0.307	0.471	0.077	0.082
SPC03	Night	N ^a	8209	3268	3604	795	542
		O/C ^b	0.279	0.199	0.228	0.519	0.747
		H/C ^b	1.266	1.2	1.195	1.584	1.668
		DBE ^{b,c}	9.198	10.438	9.916	4.283	4.161
		AI _{mod} ^{b,c}	0.223	0.184	0.31	0.101	0.061
SPC04	Day	N ^a	7707	2784	3189	1065	669
		O/C ^b	0.299	0.204	0.251	0.438	0.697
		H/C ^b	1.271	1.183	1.173	1.564	1.632
		DBE ^{b,c}	8.769	10.264	9.655	4.896	4.495
		AI _{mod} ^{b,c}	0.248	0.218	0.36	0.101	0.076
SPC04	Fog	N ^a	6086	2240	2831	738	277
		O/C ^b	0.327	0.292	0.241	0.564	0.863
		H/C ^b	1.277	1.241	1.156	1.673	1.758
		DBE ^{b,c}	7.969	8.54	9.069	3.683	3.542
		AI _{mod} ^{b,c}	0.247	0.207	0.353	0.05	0.018

^aValues represent the respective number of molecular formulas. ^bValues represent mathematical averages based on formula assignment. ^cDouble bond equivalents (DBE) and modified aromaticity index (AI_{mod}) values were obtained using equations (3.1) and (3.2) to (3.4) respectively.

Figure 3.3. Percentage of assigned molecular formulas to each of the elemental groups in the fog water and aerosol samples for -ESI (a) and +ESI (b) analyses.

SPC01: Isotopic Detection of Sulfur

Figure 3.4. Molecular formula assignments containing S were further validated by comparisons to the detection of the expected isotopic mass of the assigned formula (^{32}S vs. ^{34}S : $\Delta_{\text{mass}} = 1.9958$ Da). Molecular formulas are shown here for SPC01. The expected trend is indicated with a solid black line, where the higher intensity formulas follow this trend well. We expect some isotope dilution for lower intensity peaks, as observed here with many lower intensity formulas which are missing corresponding ^{34}S masses.

Figure 3.5. Density distribution of elemental ratios for the assigned molecular formulas of the ambient fog and aerosol samples, analyzed by -ESI (a) and +ESI (b). Oxygen-to-carbon (green) and hydrogen-to-carbon (purple) are shown. The shape width is scaled to the number of formulas with the same ratio.

Figure 3.6. Carbon number intensity trends for SPC01 daytime and nighttime (a and b) intensity trends. The difference of these intensity trends (c and d) show the enhanced intensity for formulas in daytime samples with positive values and in nighttime samples with negative values. The formulas are separated by polarity (-ESI in a and c; +ESI in b and d). Normalized intensities were used.

Figure 3.7. Carbon number intensity trends for SPC01 daytime and fog (a and b) intensity trends. The difference of these intensity trends (c and d) show the enhanced intensity for formulas in daytime samples with positive values and in fog samples with negative values. The formulas are separated by polarity (-ESI in a and c; +ESI in b and d). Normalized intensities were used.

Figure 3.8. Carbon number intensity trends for SPC01 fog and nighttime (a and b) intensity trends. The difference of these intensity trends (c and d) show the enhanced intensity for formulas in fog samples with positive values and in nighttime samples with negative values. The formulas are separated by polarity (-ESI in a and c; +ESI in b and d). Normalized intensities were used.

Figure 3.9. Carbon number intensity trends for SPC02 daytime and nighttime (a and b) intensity trends. The difference of these intensity trends (c and d) show the enhanced intensity for formulas in daytime samples with positive values and in nighttime samples with negative values. The formulas are separated by polarity (-ESI in a and c; +ESI in b and d). Normalized intensities were used.

Figure 3.10. Carbon number intensity trends for SPC02 daytime and fog (a and b) intensity trends. The difference of these intensity trends (c and d) show the enhanced intensity for formulas in daytime samples with positive values and in fog samples with negative values. The formulas are separated by polarity (-ESI in a and c; +ESI in b and d). Normalized intensities were used.

Figure 3.11. Carbon number intensity trends for SPC02 fog and nighttime (a and b) intensity trends. The difference of these intensity trends (c and d) show the enhanced intensity for formulas in fog samples with positive values and in nighttime samples with negative values. The formulas are separated by polarity (-ESI in a and c; +ESI in b and d). Normalized intensities were used.

Figure 3.12. Carbon number intensity trends for SPC03 daytime and nighttime (a and b) intensity trends. The difference of these intensity trends (c and d) show the enhanced intensity for formulas in daytime samples with positive values and in nighttime samples with negative values. The formulas are separated by polarity (-ESI in a and c; +ESI in b and d). Normalized intensities were used.

Figure 3.13. Carbon number intensity trends for SPC03 daytime and fog (a and b) intensity trends. The difference of these intensity trends (c and d) show the enhanced intensity for formulas in daytime samples with positive values and in fog samples with negative values. The formulas are separated by polarity (-ESI in a and c; +ESI in b and d). Normalized intensities were used.

Figure 3.14. Carbon number intensity trends for SPC03 fog and nighttime (a and b) intensity trends. The difference of these intensity trends (c and d) show the enhanced intensity for formulas in fog samples with positive values and in nighttime samples with negative values. The formulas are separated by polarity (-ESI in a and c; +ESI in b and d). Normalized intensities were used.

Figure 3.15. Carbon number intensity trends for SPC04 daytime and fog (a and b) intensity trends. The difference of these intensity trends (c and d) show the enhanced intensity for formulas in daytime samples with positive values and in fog samples with negative values. The formulas are separated by polarity (-ESI in a and c; +ESI in b and d). Normalized intensities were used.

3.4.1 Comparing day to night samples

Daytime aerosol is expected to initially compose the interstitial nighttime aerosol which are not activated to form fog droplets. Figs. 3.6, 3.9 and 3.12 show the day-night difference of intensity trends for SPC01, SPC02 and SPC03.

For -ESI, CHOS species are strongly enhanced in daytime and CHNO are slightly enhanced in nighttime. Overall, the distribution of C_n in -ESI analyses was roughly equal in most of the samples, except for SPC02 which was shifted towards higher C_n in the daytime sample compared to the nighttime. In this way, SPC01 also has a tailing intensity trend for CHO and CHNO species at higher C_n . Organosulfates are expected to form in the aerosol liquid water phase of deliquesced aerosol (Darer et al. 2011, Hu et al. 2011, McNeill et al. 2012). As sulfate salts and organosulfates are expected to be hygroscopic, the aerosol containing these species may have more easily been activated as fog droplets, and thus these organosulfate species would not be present in the nighttime samples.

For +ESI, CHN species are strongly enhanced at nighttime and CHNO species are enhanced in most daytime samples. In most samples, C_n was also shifted lower in nighttime samples compared to daytime (except for SPC01) suggesting some degree of chemical fragmentation occurs at night (Kroll et al. 2011). Nighttime samples have less influence from solar radiation and aerosol liquid water compared to daytime, which may lead to increased stability of the CHN species observed here. Some functionalized species related to atmospheric brown carbon are expected to form from chemical reactions in evaporating water droplets, which may lead to the formation of these CHN species in nighttime interstitial aerosol.

3.4.2 Comparing day to fog samples

Fog droplets are formed in the evening as temperatures drop and RH concurrently increases to supersaturation levels. Hygroscopic aerosol from the daytime are thus expected to act as seeds for droplet formation and thus contribute to the initial composition of fog droplets. Figs. 3.7, 3.10, 3.13 and 3.15 show the day-fog difference of intensity trends for SPC01, SPC02, SPC03 and SPC04.

For -ESI, CHOS and CHNOS species were enhanced in intensity in daytime and CHO and CHNO species were enhanced in fog. Overall the daytime aerosol is shifted to higher C_n than the fog, which is consistent with our previous observations (Brege et al. 2018) attributed to aging reactions in the aqueous phase which are expected to degrade larger C_n molecules. The CHOS and CHNOS species of the daytime aerosol were shifted to higher carbon numbers than the CHO and CHNO species of the fog. This is consistent with hydrolysis reaction in the aqueous phase which may degrade organosulfates (Darer et al. 2011, Hu et al. 2011), and lead to the enhanced intensity of CHO and CHNO species observed here.

For +ESI, there was an overall observed shift from larger to smaller C_n between daytime aerosol and fog. This is consistent with what was observed for -ESI. CHN species are enhanced in intensity in daytime compared to fog, where if these species were formed in evaporating droplets at night, they would no longer be present in fog samples. The CHN species may then remain in interstitial aerosol until daytime. Given the high influence of CHNO species in the +ESI samples (panel b in Figs. 3.6 to 3.15) it is worth mentioning the enhanced intensity of higher C_n CHO species observed in most daytime aerosol compared to fog. It is possible that higher C_n species are degraded in the fog and thus observed here with enhanced intensity in the daytime aerosol.

3.4.3 Comparing night to fog samples

As stated previously, fog droplets are formed from hygroscopic aerosol in supersaturation conditions. Aerosol particles which do not form fog droplets remain in the interstitial space and would be expected to have different molecular compositions than the fog, though both nighttime aerosol and fog would be expected to have influences from daytime aerosol compositions. Figs. 3.8, 3.11 and 3.14 show the night-fog difference of intensity trends for SPC01, SPC02 and SPC03.

For -ESI, nighttime aerosol was enhanced in intensity at higher C_n compared to fog. There was some influence from CHOS species at higher C_n which may be related to influences from daytime aerosol. As described above for daytime-nighttime comparisons, CHOS and CHNOS species were not expected to remain in the nighttime interstitial aerosol during fog formation, meaning that any minor enhancement of these species in fog may skew the distributions observed here. As these S-containing species may degrade in fog (Darer et al. 2011, Hu et al. 2011), the enhanced intensity observed for fog in nighttime-fog comparisons may indicate that some other factors, such as LWC or pH, also play a role in these degradation processes.

For +ESI, CHN species were enhanced in intensity in nighttime aerosol, and CHNO species were enhanced in fog. This is consistent with what was observed for daytime-nighttime comparisons above. Fog once again was distributed towards lower C_n with most nighttime aerosol samples shifted to higher C_n . The enhanced intensity of CHN species in nighttime aerosol compared to fog is consistent with what was observed for daytime-nighttime comparisons. If these CHN species are related to brown carbon, it is possible that they are formed in evaporating droplets (Lee et al. 2013, Hawkins et al. 2018) which would remove them from fog.

3.4.4 Pyrazine chromophores

Chromophoric pyrazine based molecular species have been reported from laboratory experiments of aqueous phase reactions of ammonia and methylglyoxal laboratory experiments (Hawkins et al. 2018). The formation mechanism for these pyrazines is described via two steps; ammonia replaces the aldehyde of methyl glyoxal to form a primary amine, then two units can undergo dehydration to combine and form 2,5-dimethylpyrazine

(Hawkins et al. 2018). Other types of aqueous phase laboratory experiments have involved evaporation as a step in the formation of laboratory light absorbing products, which is expected to remove water from the system and drive reaction equilibria forward (De Haan et al. 2011, Nguyen et al. 2012, Lee et al. 2013). Hawkins et al. (2018) report 18 molecular species (see Table 3.4) suspected to be pyrazine ($C_4H_4N_2$) based chromophoric molecular species from the methyl glyoxal-ammonium sulfate reactions. We observe 12 to 14 of the same molecular formulas in the ambient fog and aerosol samples, particularly N_2 and CHNO species (Table 3.4). Fig. 3.16 shows the nitrogen intensity trends observed for species with estimated aromatic and condensed aromatic character ($AI_{mod} > 0.5$) (Koch and Dittmar 2006, 2016) in the ambient fog and aerosol samples. Here we observe an enhanced intensity of the estimated aromatic and condensed aromatic species with $N \geq 1$ compared to N_0 in almost all samples. There were a significant number of estimated aromatic and condensed aromatic species ($AI > 0.5$) with N_2 in our samples which may be related to the pyrazine-based chromophores described in Hawkins et al. (2018). We expect significant influence of ammonia from agriculture in the Po Valley, and similar reactions with aldehyde groups present in atmospheric organics from biomass burning (Ervens et al. 2013, Nguyen et al. 2013, Galloway et al. 2014) may lead to the formation of chromophoric species by similar processes.

Laboratory experiments similar to those described in Hawkins et al. (2018) were conducted in small quantities (150 μ L) of equal volumes 1 M methyl glyoxal and 1 M ammonium sulfate. These reactions were adjusted to pH 2 to facilitate the expected acid catalyzed formation of the pyrazine-like compounds. The reaction was left open to the air to allow for evaporation which was expected to further facilitate the reaction. The sample was compared to an experimental blank which did not contain ammonium sulfate. The samples were re-dissolved in acetonitrile and analyzed by OTE-MS in a similar manner as the SPC samples, though atmospheric pressure photoionization (APPI) in the positive polarity mode was used in an attempt to fully characterize the expected light absorbing species. While the ionization modes are different between this laboratory sample and the ambient SPC samples, our related work (see chapter 3) suggests that there is a strong overlap between the molecular formulas observed between +APPI and +ESI (~65 to 70%).

The results of the OTE-MS analysis of the methyl glyoxal-ammonium sulfate laboratory sample are provided in Fig. 3.17. Overall ~1400 molecular formulas were assigned to the sample, including all of the molecular formulas for the 18 suspected pyrazine-based chromophores described in Hawkins et al. (2018) (marked with diamonds in the mass spectra of Fig. 3.17a). A majority of the formulas were CHO and CHNO followed by CHN species. A strong enhanced intensity for C_6 , C_9 , C_{12} and C_{15} species was observed. A majority of the intensity is attributed to formulas with AI_{mod} that suggest olefinics followed by aromatics, with minor contributions from aliphatic and condensed aromatics. We note that the light absorbing components may be present in low abundance in the sample, and that it is not certain which molecular formulas contribute to light absorption.

Table 3.4. Molecular formulas for pyrazine-based chromophores identified in Hawkins et al. (2018) observed in SPC 2015 ambient samples.

	SPC01			SPC02			SPC03			SPC04	
Molecular formula (Hawkins et al. 2018)	Day	Night	Fog	Day	Night	Fog	Day	Night	Fog	Day	Fog
C ₄ H ₄ N ₂											
C ₄ H ₆ N ₂											
C ₅ H ₈ N ₂											
C ₆ H ₈ N ₂	X	X	X	X	X	X	X	X	X	X	X
C ₆ H ₈ N ₂ O	X	X	X	X	X	X	X	X	X	X	X
C ₆ H ₉ NO ₃	X		X	X		X		X	X		X
C ₉ H ₉ NO ₃	X	X	X	X	X	X	X	X	X	X	X
C ₉ H ₁₁ N ₃											
C ₉ H ₁₂ N ₂ O ₂	X	X	X	X	X	X	X	X	X	X	X
C ₉ H ₁₂ N ₂ O ₃	X	X	X	X	X	X	X	X	X	X	X
C ₉ H ₁₃ NO ₅	X	X	X	X		X	X		X	X	X
C ₁₂ H ₁₁ NO ₄	X	X	X	X	X	X	X	X	X	X	X
C ₁₂ H ₁₃ N ₃ O ₂	X	X	X	X	X	X	X	X	X	X	X
C ₁₂ H ₁₄ N ₂ O ₃	X	X	X	X	X	X	X	X	X	X	X
C ₁₂ H ₁₄ N ₂ O ₄	X	X	X	X	X	X	X	X	X	X	X
C ₁₂ H ₁₆ N ₂ O ₄	X	X	X	X	X	X	X	X	X	X	X
C ₁₂ H ₁₇ NO ₇	X	X	X	X	X	X	X	X	X	X	X
C ₁₅ H ₁₆ N ₂ O ₄	X	X	X	X	X	X	X	X	X	X	X

Figure 3.16. Nitrogen number intensity trends for molecular formulas with high modified aromaticity index values ($AI_{mod} > 0.5$) implying aromatic and condensed aromatic character (Koch and Dittmar 2006, 2016). Panels are separated by SPC01 (a), SPC02 (b), SPC03 (c) and SPC04 (d). Intensities were summed using normalized intensities.

We matched the molecular formulas of the ambient SPC fog and aerosol samples to this laboratory sample, resulting in a subset of ~800 to 1100 (~16 to 23%) molecular formulas from each of the ambient samples. The overlapping molecular formulas from the laboratory and ambient samples are presented in Figs. 3.18, 3.19 and 3.20. We observe that many CHNO formulas were in common between the laboratory and ambient samples, with significant influence from species with $N > 1$. We also observe that the aromatic and condensed aromatic species had higher intensity in the ambient samples compared to the laboratory sample. The enhanced intensity trends observed for C_6 , C_9 , C_{12} and C_{15} species in the laboratory sample are not present in the ambient samples, except for C_6 . This may mean that if pyrazine compounds are formed in ambient fog and aerosol they may quickly be further functionalized into other species. These trends may also simply be masked by other species present in the ambient samples. We observe a strong influence of brown carbon observed in these samples (Table 3.1), significant estimated aromatic molecular characteristics and many overlapping molecular formulas in laboratory and ambient samples. Based on these observations, we suggest the possibility of pyrazine-based chromophoric species in ambient fog and aerosol, in addition to a host of other molecular species.

3.5 Conclusions

In the comprehensive analysis of ambient fog and aerosol samples described here, we observe extraordinary molecular level diversity. The molecular species of daytime, nighttime and fog water followed predictable trends from our previous work (Brege et al. 2018), but have thus far been unobserved for chronologically related samples. The formation of organosulfates in aerosol liquid water has been observed previously, however the CHOS molecular species in the daytime samples were observed here with much higher intensities and in greater numbers than our previous study (Brege et al. 2018). The fog processing of biomass burning organic aerosol has been shown previously to enhance aerosol light absorption (Gilardoni et al. 2016) and the trends described for here for N-containing chemical species with high estimated aromatic character may help to explain those observations. Furthermore, these N-containing molecular formulas with high estimated aromatic character may also be related to pyrazine, imine and imidazole chromophores. Pyrazines are expected to form in a two-step chemical reaction between ammonia and methyl glyoxal (Hawkins et al. 2018). If these reactions can also take place for other organics with aldehyde functional groups, the observation of these chemical species in these samples implies that this reaction mechanism is viable for fog water and aerosol liquid water in the region.

Figure 3.17. Summary of the molecular formula assignment of a laboratory aqueous phase secondary organic aerosol (aqSOA) methyl glyoxal-ammonium sulfate mixture, which is suspected to produce pyrazine-based chromophores. The reconstructed mass spectra (a) carbon number intensity trends (b) and nitrogen number intensity trends (c) are shown. The modified aromaticity index (AI_{mod}) is described at the end of section 3.3.

Figure 3.18. Summary of the molecular formulas which were observed in both the daytime ambient aerosol samples from Po Valley and the laboratory aqueous phase secondary organic aerosol (aqSOA) methyl glyoxal-ammonium sulfate mixture, which is suspected to produce pyrazine-based chromophores. Carbon number (upper panels) and nitrogen number (lower panels) trends are shown for the molecular formulas which were present in both data sets.

Figure 3.19. Summary of the molecular formulas which were observed in both the nighttime ambient aerosol samples from Po Valley and the laboratory aqueous phase secondary organic aerosol (aqSOA) methyl glyoxal-ammonium sulfate mixture, which is suspected to produce pyrazine-based chromophores. Carbon number (upper panels) and nitrogen number (lower panels) trends are shown for the molecular formulas which were present in both data sets.

Figure 3.20. Summary of the molecular formulas which were observed in both the ambient fog water samples from Po Valley and the laboratory aqueous phase secondary organic aerosol (aqSOA) methyl glyoxal-ammonium sulfate mixture, which is suspected to produce pyrazine-based chromophores. Carbon number (upper panels) and nitrogen number (lower panels) trends are shown for the molecular formulas which were present in both data sets.

3.6 Associated content

3.6.1 Data availability

The final completed data set for this work is intended to be made available alongside the published results once the manuscript is completed and submitted.

3.6.2 Acknowledgements

We thank Dr. Lelia Hawkins for her generous provision of ultrahigh resolution mass spectrometry data for the laboratory methyl glyoxal-ammonium sulfate aqueous phase secondary organic aerosol experiments for comparison to the ambient samples of this data set.

Many supporting measurements and calculations for this chapter were prepared by our collaborators at the Institute of Atmospheric Sciences and Climate (Italian National Research Council). Specifically, Dr. Maria Cristina Facchini, Dr. Stefano Decesari, Dr. Stefania Gilardoni, Dr. Marco Paglione and Julija Pauraitė have each contributed ambient samples for analysis or have provided processed data from field measurements. In the publication of this chapter they are intended to be listed as co-authors.

4 Extreme molecular complexity resulting in a continuum of carbonaceous species in biomass burning tar balls³

4.1 Abstract

Biomass burning emits numerous carbonaceous particles into the atmosphere. Spherical organic aerosol particles, commonly known as tar balls, represent the most abundant particles generated by biomass burning. However, the detailed molecular-level composition of tar balls is largely unknown. Two atmospheric aerosol samples, heavily influenced by tar balls, were analyzed using ultrahigh resolution Orbitrap Elite mass spectrometry with four complementary ionization modes. The molecular formulas were primarily composed of CHO and CHNO species, with the number of unambiguous molecular formulas observed per sample being in excess of 10,000. The estimated saturation mass concentrations and relative humidity dependent glass transition temperatures for the CHO molecular formulas were consistent with an overall low volatility and solid morphology expected for tar balls. The measurement of room-temperature evaporation kinetics showed that these particles are composed of ~90% non-volatile species. The degree of molecular complexity detected here signifies a continuum of carbonaceous species, observed from C₃ to C₄₅ and with continuous ranges of oxygenation and hydrogen saturation for each C_n. Within this continuum we observed approximately 24% of molecular formulas had high estimated aromatic character, which could be related to chemical species with negative health effects and species with light absorption properties. The continuum of carbonaceous species observed here for samples influenced mainly by tar balls has strong implications for the molecular characterization of atmospheric organic matter. The level of complexity detected here cannot be ignored in future aerosol analyses. We demonstrate that multiple analytical methods may be required to suitably interpret this complexity on a molecular level.

4.2 Introduction

Wildfires in North America have steadily increased in frequency and duration attributed to dryer conditions brought on by increased global temperatures (Westerling et al. 2006, Barbero et al. 2015, Gergel et al. 2017). The extreme temperatures of wildfires can create convective forces capable of elevating emissions into the free troposphere and the stratosphere, facilitating long range transport (Damoah et al. 2004, Koppmann et al. 2005, Yu et al. 2019). In addition to wildfires, residential biomass burning for heating and cooking are significant source of carbon compounds to the atmosphere (Simoneit 2002, Fleming et al. 2018, Fleming et al. 2018), where it is estimated that 75% of the world

³As stated in the Preface, the material contained in chapter 4 was prepared with the intent of submission for publication in the journal *Environmental Science & Technology*. As of 9-Dec 2019 the manuscript has not yet been submitted.

population uses wood as a fuel source (Koppmann et al. 2005). Vapor phase emissions are transformed by atmospheric oxidants into lower volatility species, which can undergo gas to particle phase conversion; these processes can form both new particles and organic coatings on existing aerosol (Kroll and Seinfeld 2008, Zhang 2010). Collectively, aerosol particles impose significant risks to human health and global climate (Koppmann et al. 2005, WHO 2016, IPCC 2018). Biomass burning emissions include mutagenic nitrophenols (Harrison et al. 2005) and carcinogenic polycyclic aromatic hydrocarbons (PAHs) (Farmer 2003, Xue and Warshawsky 2005) while the catalytic generation of reactive oxygen species has been associated in particular with biomass burning aerosol and humic-like substances (Lin and Yu 2011, Verma et al. 2015). Prolonged exposure to fine particulate matter is associated with cardiovascular and pulmonary diseases (WHO 2016). Light absorbing “black carbon” and “brown carbon” contribute to atmospheric warming and are associated with biomass burning emissions (Andreae and Gelencser 2006, Saleh et al. 2018). The potential energy absorption for black carbon is approximately 5 orders of magnitude greater than an equivalent mass of CO₂, with brown carbon having roughly 25% the absorbing potential of black carbon (Laskin et al. 2015). Moderating anthropogenic absorbing aerosol in the atmosphere may help to mitigate the effects of climate change on a time scale as little as 5 years (Jacobson 2002).

Organic spherical particles known as “tar balls”, are a morphologically defined sub-component of atmospheric aerosol associated with the smoldering phase of biomass burning (Reid et al. 2005). Tar balls are approximately 100 to 300 nm in size, glassy-like and highly viscous, with no observable condensation nuclei or concentric layers (Reid et al. 2005, Reid et al. 2018, Sedlacek et al. 2018). Tar balls can resist the high-powered electron beam of microscopy techniques where other organic particles are partially or fully vaporized (Posfai et al. 2004). In wildfires, tar balls have been observed in much greater numbers compared to soot particles, at an abundance of approximately 10:1 (Chakrabarty et al. 2010, China et al. 2013, Sedlacek et al. 2018). The increased detection of tar balls directly downwind of biomass burning plumes indicates they are not directly emitted, but somewhat aged particles formed over a matter of hours (Sedlacek et al. 2018, Adachi et al. 2019). Tar balls contain light absorbing brown carbon (Chakrabarty et al. 2010) and optically “dark” and “light” tar balls have been observed (China et al. 2013). A wide range of wavelength dependence to light absorption has been described in the literature (Alexander et al. 2008, Chakrabarty et al. 2010, Bhandari 2018) implying that the molecular makeup of tar balls between different sample sources may be inconsistent.

The chemical characterization of tar balls is an active area of research given their high abundance in biomass burning smoke and light absorption properties. Due to the intricacies of combustion for solid fuels, such as wood in biomass burning, the complexity of byproducts is greatly enhanced compared to liquid and gas phase fuels (Wielgosinski 2012). The biopolymer lignin constitutes approximately 30% of wood tissue and many wood burning products have chemical structures related to the phenolic monomers which make up lignin (Simoneit 2002, Reid et al. 2005). The diversity of chemical products is highly dependent on the combustion conditions of the fire, including fuel type, moisture content, fire size, oxygen availability and the progression of flaming versus smoldering

phases (Koppmann et al. 2005, Reid et al. 2005, Wielgosiński 2012). Broad classifications of chemical species are associated with biomass burning emissions, including alcohols, aldehydes, ketones, carboxylic acids, esters, ethers, furanes, sugars and anhydrosugars, lignans, nitro-aromatics and PAHs (Andreae and Merlet 2001, Simoneit 2002, Koppmann et al. 2005, Mazzoleni et al. 2007, Samburova et al. 2016). Particularly for tar balls, methoxy, ketone and carboxylic acid functional groups are expected with influence from aromatic and phenolic structures (Adachi et al. 2019, Li et al. 2019). Transmission electron microscopy and energy-dispersive X-ray spectroscopy analyses describe elemental compositions of primarily carbon and oxygen, in carbon to oxygen ratios of 1 to 10 (Chakrabarty et al. 2010, China et al. 2013). A recent study by Adachi et al. (2019) describes the importance of organic nitrogen compounds in the formation of tar balls, expected to be in the form of organic nitrate and nitro-aromatic species.

The molecular level characterization of tar balls has the same analytical challenges as any other complex mixture of organics. In recent years analytical approaches using ultrahigh-resolution mass spectrometry have become preferable in analyzing these types of mixtures, because of their ability to efficiently resolve the monoisotopic masses of chemical species with the same nominal mass (Marshall 2000, Zubarev and Makarov 2013). These “bottom-up” style analyses for complex mixtures are often implemented in that analytical standards for comparison are nonexistent. “Soft-ionization” techniques like electrospray ionization (ESI) produce charged analytes with little to no fragmentation (Konermann et al. 2013) and when combined with ultra-high resolution techniques ESI has been used successfully to unveil the molecular complexity of ambient organic aerosol on multiple occasions (Altieri et al. 2009, Schmitt-Kopplin et al. 2010, Lin et al. 2012, Zhao et al. 2013, Wozniak et al. 2014, Dzepina et al. 2015, Willoughby et al. 2016, Cook et al. 2017). ESI is selective to molecules with polar functional groups capable of easily accepting or removing a proton (Konermann et al. 2013) which can unfortunately lead to the suppressed ionization of molecules without polar groups through charge competition. Atmospheric pressure photoionization (APPI) is an alternate soft ionization method which is more sensitive to low polarity and aromatic molecules with respect to ESI (Kauppila et al. 2002). In this study, the acetonitrile-soluble extracts of two biomass burning samples heavily influenced by tar ball morphology were analyzed using ultrahigh-resolution Orbitrap Elite mass spectrometry (OTE-MS) to determine their molecular formula composition. We have applied an extensive molecular level analysis using the negative and positive polarities of ESI and APPI, hereafter abbreviated as -APPI, +APPI, -ESI, and +ESI. The observed result is an unprecedented degree of molecular complexity. We will show how this combination of ionization techniques provides a more comprehensive view of the complex mixture of organic matter in atmospheric aerosol and demonstrates a continuum of carbonaceous chemical species in organic aerosol.

NOAA HYSPLIT MODEL
Backward trajectories ending at 1800 UTC 05 Sep 17
GDAS Meteorological Data

Figure 4.1. HYSPLIT 48 h back trajectory model for Richland, WA on 5-Sep 2017. Wildfires burned in the central and northern areas of the state, in the approximate areas of the back trajectory.

4.3 Methods

In early September 2017 several wildfires burned in the Pacific Northwest, including the Uno Peak and Sawmill Creek fires in central Washington (NWCC 2018). Two aerosol filter samples were collected at Pacific Northwest National Lab (Richland, WA) on 5-Sep 2017 and 6-Sep 2017 and will be referred to here as BB05 and BB06, respectively. HYSPLIT back trajectories show influence from northern Washington state for the preceding 48 hours of these samples (Fig. 4.1) incorporating the approximate locations of the fires. The samples were collected on 90 mm PTFE filters by pulling air at 80 L min^{-1} for 6 hours (a sample blank was also prepared). Single particle analysis using scanning electron microscopy (SEM) of the filter samples suggest that tar balls were the most abundant particle type collected on the sample filters (~85 to 95% number fraction) (Fig. 4.2). Similarly, real-time characterization of individual aerosol particles using a single particle

Figure 4.2. Scanning electron microscopy image of the aerosol filter sample collected 5-Sep 2017 (BB05). Tar balls were the most abundant particle type collected on the filter sample (~85-95% number fraction).

mass spectrometer (Zelenyuk et al. 2015) indicated that the vast majority (>95%) of sampled aerosol particles were spherical biomass burning particles, i.e. tar balls. Additionally, scanning transmission X-ray microscopy (STXM) and near-edge X-ray absorption fine structure (NEXAFS) spectra of the particles (Fig. 4.3) agree with previous results described for tar balls (Tivanski et al. 2007). In addition to filter sampling for offline analysis, online particle characterization was conducted using a scanning mobility particle sizer (SMPS) and single particle mass spectrometer (miniSPRAT). Real time measurements of particle size distributions, density, shape, mass loadings, single particle mass spectra, and volatility were obtained (Vaden et al. 2011). The average size distributions of BB05 and BB06 are shown in (Fig. 4.4) and using the measured density of 1.4 g cm^{-3} yields an average aerosol mass loading of $170 \pm 30 \text{ } \mu\text{g m}^{-3}$.

During preparation for sample analysis by OTE-MS, high purity analytical grade solvents were used in every step. A portion of the aerosol filter was extracted in acetonitrile (ACN), a polar organic aprotic solvent, using a low speed orbital shake table for 90 minutes at 60 RPM. The extract was filtered using a pre-conditioned 0.2 μm PTFE membrane syringe filter to remove insoluble matter. ACN was used in an effort to extract the low-polarity organics which are not water-soluble and are expected to be present in tar balls (Tfaily et al. 2015), while many of the organics which are water-soluble are also expected to be extracted by ACN. The results described here are thus limited to those chemical species which are extracted by ACN.

Figure 4.3. Near-edge X-ray absorption fine structure (NEXAFS) spectra of BB05 filter collected particles. The absorption peaks observed at 285.1, 286.7 and 288.5 agree with previous results described for tar balls and correspond to aromatic/phenolic, ketone/carbonyl and carboxylic acid functionalities (Tivanski et al. 2007).

Figure 4.4. Real time measurements of particle size distributions for BB05 and BB06 measured by scanning mobility particle sizer. Particle counts measured with a resolution of finite size bins can skew results between instruments with different bins. The normalized particle concentration ($dN/d\log D_p$) is independent of bin width, where $d\log D_p$ is calculated as the difference in the log of the lower bin boundary from the log of the upper boundary for each channel and dN is the number of respective particles in that range (Hinds 1999).

The samples were analyzed using OTE-MS in the Chemical Advanced Resolution Methods (ChARM) Laboratory at Michigan Technological University (Houghton, Michigan, USA). Samples were directly infused into the instrument using a 250 μL syringe pump at 5 to 30 $\mu\text{L min}^{-1}$. The analyses used APPI and ESI in both the positive and negative polarities, for a total of four ionization methods. Briefly, APPI uses photolytic ionization via a UV lamp, and is expected to ionize less polar and more aromatic species. ESI uses a charged capillary to create a fine spray of evaporating solvent, while also applying a charge, and is expected to ionize more polar organics. For +APPI and -APPI, the samples were diluted by a factor of 2 and 1% toluene was added as a dopant for the UV lamp (Kauppila et al. 2002). The samples were diluted by a factor of 4 for -ESI and by a factor of 8 for +ESI to optimize the ion signal. Data was collected for 200 scans over the range of m/z 100 to 800 and at a resolution of 240,000 (defined at m/z 400) with spectrum averaging.

Mass lists were generated from an average of the 200 scans using Xcalibur software (Thermo Scientific v. 3.0). The newly developed MFAssignR (Schum et al. 2018) was used to estimate the noise threshold, recalibrate the spectra and assign molecular formulas to the measured masses. A series of custom R scripts (R v. 3.5.1 and RStudio v. 1.1.463) were used to perform blank subtraction and authenticate data quality. The detailed mechanisms for molecular formula assignment using MFAssignR are beyond the scope of this paper, and are described elsewhere (Schum et al. 2018) while a detailed synopsis of formula assignment using MFAssignR for this data set is provided in section 4.6. Briefly, molecular formulas were assigned allowing for carbon, hydrogen, oxygen, nitrogen and sulfur elements with limits placed on oxygen (0 to 30), nitrogen (0 to 3) and sulfur (0 to 1). Up to one sodium atom was allowed for data collected in +ESI to allow for positively charged sodium adducts in addition to hydrogen adducts. Formulas were further validated by a series of rigorous quality assurance steps as in our previous work (Brege et al. 2018, Schum et al. 2018), which is outlined in section 4.6. The reconstructed mass spectra of BB05 using our comprehensive multi-ionization OTE-MS analysis are shown in Fig. 4.5. In each spectrum we observed ~4000 to 7000 formulas depending on ionization mode and we observed ~10,000 formulas in each sample once accounting for formulas which were detected in multiple ionization modes (see section 4.6.3). Overall, there was good agreement between the two samples with respect to the number of molecular formulas detected with each ionization method (Figs. 4.6 and 4.7). The combined samples identified ~12,000 total molecular formulas for biomass combustion tar balls.

Figure 4.5. Reconstructed mass spectra for BB05 as analyzed by multiple ionization methods. Only ion peaks which were assigned a valid molecular formula are shown.

Figure 4.6. The distribution of all 10,887 molecular formulas of BB05 in van Krevelen space (center) with the density distribution of the formulas by O/C and H/C (top and right respectively) in order to visualize the magnitude of overlapping points in the center panel. Points (and bars) are colored by the method of detection, whether a formula was exclusive to one of the four ionization modes (red, purple, blue, gold) detected in all four modes (green) or some other combination of multiple ionization modes (gray). Dashed lines represent $H/C = 1.2$ (vertical), $O/C = 0.6$ (horizontal) and $OS_C = 0$ (diagonal) (Tu et al. 2016).

Figure 4.7. The distribution of all 10,553 molecular formulas of BB06 in van Krevelen space (center) with the density distribution of the formulas by O/C and H/C (top and right respectively) in order to visualize the magnitude of overlapping points in the center panel. Points (and bars) are colored by the method of detection, whether a formula was exclusive to one of the four ionization modes (red, purple, blue, gold) detected in all four modes (green) or some other combination of multiple ionization modes (gray). Dashed lines represent $H/C = 1.2$ (vertical), $O/C = 0.6$ (horizontal) and $O/C = H/C$ (diagonal) (Tu et al. 2016).

The accepted molecular formulas were classified into five elemental groups: "CH", "CHO", "CHN", "CHNO" and "CHOS", based on the respective inclusion of carbon, hydrogen, oxygen, nitrogen and sulfur elements in the formula. We note that each identified molecular formula likely represents multiple different molecular isomers at extremely small concentrations (Zark et al. 2017, Hawkes et al. 2018) and complete chemical structural elucidation is impossible without chromatographic separation and additional tandem mass spectrometry (MS²) analysis. However, the elemental composition of the molecular formulas can be used to estimate chemical properties through calculations, such as the number of double bonds or ring structures present, or double bond equivalents (DBE). The modified aromaticity index (AI_{mod}) was used to estimate the number of carbon-carbon unsaturated bonds (aromatic character) and was calculated as described by Koch and Dittmar (2006, 2016). For condensed aromatics, we also required formulas to have 8 or more carbon atoms to avoid the misclassification of formulas which do not have a reasonable number of carbon atoms to form multiple ring structures. Molecular formulas were separated into four groups based on their calculated AI_{mod} values that are defined in Table 4.1. The average oxidation state of carbon (OS_C) was estimated as described by Kroll et al. (2011). The calculations for DBE, AI_{mod} and OS_C are described in detail in section 4.6 with equations (4.3) to (4.8).

The volatility of the molecular formulas, specifically the saturation mass concentration (C₀ in units of µg m⁻³) was estimated using a model developed by Li et al. (2016). Molecular formulas were separated into five groups based on their calculated C₀ values that are defined in Table 4.1. In addition to the calculated volatility of the molecular formulas, the volatility of overall particles was determined using measurements of room-temperature evaporation kinetics of size-selected particles, described in detail elsewhere (Vaden et al. 2011). The relative humidity dependent glass transition temperature (T_{g,RH}) was estimated for CHO formulas based on the work of DeRieux et al. (2018). The ratio of T_{g,RH} to the average ambient temperature (T_{amb}) during sample collection (Table 4.2) can be used to estimate the phase state ratio of the molecular formulas as defined in DeRieux et al. (2018). The molecular formulas were separated into three groups based on this calculated ratio that are defined in Table 4.1. The equations and parameters used to calculate C₀ and T_{g,RH} are described in detail in section 4.6 with equations (4.9) to (4.13).

4.4 Results and Discussion

4.4.1 Molecular composition of tar balls

In our comprehensive OTE-MS analysis of the ACN-soluble fraction of ambient aerosol heavily influenced by tar ball morphology, we observe that most molecular formulas in both samples are CHO or CHNO species with minor contributions from the CH, CHN and CHOS groups (Fig. 4.8). The significant number of CHNO species observed here (~50% of formulas per sample) is consistent with a recent STXM chemical characterization of tar balls which describe organonitrates (Adachi et al. 2019). A summary of BB05 and BB06 molecular formulas by elemental group is provided in Tables 4.3 and 4.4 respectively. An

overall continuum of carbonaceous chemical species is observed from the vast number of molecular formulas detected. The molecular formulas are continuous from C₃ to C₄₅ with consecutive ranges of oxygen numbers and degrees of unsaturation (represented by DBE values) for each C_n. These trends can be observed in the DBE vs. carbon number space (Figs. 4.9 and 4.10) accompanied by distribution histograms of DBE and carbon number on the respective axes.

The calculated AI_{mod} values for the molecular formulas trend well with the observed DBE and carbon number trends in this space. Formulas with low carbon numbers (~C₃ to C₁₀) and low DBE (~0 to 5) are predicted to be mostly aliphatic and olefinic species, while formulas with high DBE (~15 to 20) and moderate carbon numbers (~C₂₅ to C₃₅) are predicted to be mostly aromatic and condensed aromatic species. Oxygenation was the most intense for formulas between ~C₂₀ to C₃₀, while overall the upper limit in oxygenation increased with carbon number. Formulas with C₃₀ to C₄₅ had consistently higher oxygen number distributions compared to formulas with C₃ to C₁₀. This enhanced oxygenation should be expected for larger molecules as there are more opportunities for the addition of reactive species, especially considering we see more carbon-carbon bond saturation (represented by AI_{mod}) in formulas with higher carbon numbers. Given the overall low level of observed oxidation in these molecular formulas (average O/C of ~0.36) the presence of higher molecular weight oxygenated species is a point of interest. The area of highest molecular formula density in this space is located from approximately C₁₄ to C₂₆ and DBE 8 to 14. These molecular formulas trend well with the overall observed characterization for the samples; the formulas are majorly CHNO followed closely by CHO species, with AI_{mod} values that predict mostly olefinic accompanied by a smaller number of aromatic species. We observe low numbers of formulas at extremely low DBE (0 to 3) especially for carbon numbers above C₁₀. As these species would be expected to be extremely aliphatic and low polarity, we do not expect them to be extracted by ACN. At higher carbon numbers (C₃₅ to C₄₅) and higher DBE values (15 to 20) we also observe few molecular formulas, which we expect would be condensed aromatic species and similarly insoluble in ACN.

Table 4.1. Defined groups for molecular formulas estimating aromaticity, volatility and phase state.

Modified aromaticity index (AI_{mod}) (Koch and Dittmar 2006, 2016)	
Aliphatic	AI _{mod} = 0
Olefinic	0 < AI _{mod} ≤ 0.5
Aromatic	0.5 < AI _{mod} ≤ 0.67
Condensed aromatic	AI _{mod} > 0.67 and C _n ≥ C ₈
Estimated saturation mass concentration (C₀) (µg m⁻³) (Li et al. 2016)	
Volatile organic compounds (VOC)	C ₀ > 3 x 10 ⁶
Intermediate volatility organic compounds (IVOC)	300 < C ₀ ≤ 3 x 10 ⁶
Semi-volatile organic compounds (SVOC)	0.3 < C ₀ ≤ 300
Low-volatile organic compounds (LVOC)	3 x 10 ⁻⁴ < C ₀ ≤ 0.3
Extremely low-volatile organic compounds (ELVOC)	C ₀ ≤ 3 x 10 ⁻⁴
Estimated phase state ratio (PSR) (estimated relative humidity dependent glass transition temperature to ambient sample collection temperature) (T_{g,RH}/T_{amb}) (DeRieux et al. 2018)	
Solid	PSR ≥ 1
Semi-solid	0.8 ≤ PSR < 1
Liquid	PSR < 0.8

Table 4.2. Meteorological conditions at the sampling location during sample collection.

Date	Max. temp. (°C)	Min. temp. (°C)	Avg. temp. (°C)	Prevailing wind direction	Avg. wind speed (km h ⁻¹)	Peak gust wind speed (km h ⁻¹)	Avg. relative humidity
2019/9/5	33	22	27	NW	15	32	30.5%
2019/9/6	32	18	25	NW	8	20	33.8%

Table 4.3. BB05 molecular formula assignment summary.

Negative mode	Source	APPI				ESI			
	Group	All	CHO	CHNO		All	CHO	CHNO	CHOS
	DBE ^a	10.28	10.31	10.23		9.19	9.53	9.35	2.29
	O/C ^a	0.38	0.36	0.40		0.51	0.46	0.54	0.85
	H/C ^a	1.07	1.12	1.02		1.18	1.18	1.13	1.71
	C _n ^a	19.42	20.93	17.46		19.41	20.91	18.18	8.66
	O _n ^a	6.78	6.98	6.52		9.27	9.31	9.46	6.53
	N ^b	4222	2387	1835		5453	3146	2107	200
Positive mode	Source	APPI				ESI			
	Group	All	CHO	CHNO	CH	All	CHO	CHNO	CHN
	DBE ^a	10.13	10.28	9.99	4.64	8.85	9.04	8.73	5.97
	O/C ^a	0.29	0.27	0.32	0.00	0.29	0.34	0.25	0.00
	H/C ^a	1.11	1.14	1.07	1.31	1.31	1.30	1.32	1.25
	C _n ^a	19.89	21.79	17.78	10.73	22.19	24.03	20.76	9.79
	O _n ^a	5.44	5.48	5.43	0.00	6.21	7.73	4.97	0.00
	N ^b	4669	2482	2176	11	7152	3334	3755	63

^aValues represent mathematical averages based on formula assignment. ^bValues represent the respective number of molecular formulas.

Table 4.4. BB06 molecular formula assignment summary.

Negative mode	Source	APPI			ESI			
	Group	All	CHO	CHNO	All	CHO	CHNO	CHOS
	DBE ^a	9.58	9.02	10.24	9.11	9.39	9.29	2.31
	O/C ^a	0.38	0.37	0.38	0.50	0.46	0.54	0.90
	H/C ^a	1.13	1.21	1.04	1.19	1.19	1.14	1.69
	C _n ^a	19.17	20.16	18.01	19.31	20.82	18.16	8.14
	O _n ^a	6.70	6.93	6.44	9.19	9.20	9.38	6.59
	N ^b	4135	2222	1913	5415	3033	2201	181
Positive mode	Source	APPI			ESI			
	Group	All	CHO	CHNO	All	CHO	CHNO	CHN
	DBE ^a	9.44	9.51	9.35	8.42	8.69	8.20	5.72
	O/C ^a	0.30	0.28	0.33	0.31	0.34	0.27	0.00
	H/C ^a	1.15	1.18	1.11	1.34	1.34	1.34	1.28
	C _n ^a	19.21	21.14	17.03	22.05	24.32	20.01	9.78
	O _n ^a	5.37	5.41	5.32	6.49	7.81	5.30	0.00
	N ^b	4730	2514	2216	7213	3564	3584	65

^aValues represent mathematical averages based on formula assignment. ^bValues represent the respective number of molecular formulas.

Figure 4.8. Percent composition of molecular formulas in BB05 and BB06 by elemental groups.

Figure 4.9. Visualization of the carbon continuum of molecular formulas observed in BB05. The distribution of molecular formulas is shown in the double bond equivalents (DBE) vs. carbon number space (center panel) color scaled to the density of molecular formulas. The density distribution of the formulas by carbon number and DBE (top and right panels respectively) are subsequently color scaled to the number of oxygen atoms in each formula and to modified aromaticity index (AI_{mod}) (Koch and Dittmar 2006, 2016).

Figure 4.10. Visualization of the carbon continuum of molecular formulas observed in BB06. The distribution of molecular formulas is shown in the double bond equivalents (DBE) vs. carbon number space (center panel) color scaled to the density of molecular formulas. The density distribution of the formulas by carbon number and DBE (top and right panels respectively) are subsequently color scaled to the number of oxygen atoms in each formula and to modified aromaticity index (AI_{mod}) (Koch and Dittmar 2006, 2016).

Our results show strong overlap of molecular formulas across multiple ionization modes. These results differ from previously published work analyzing atmospheric aerosol samples via multiple ionization methods (Lin et al. 2018) and the disparity may be attributed to the differences in sample extraction solvent or possibly the chemical nature of the different samples between the two studies. Nevertheless, approximately 16% of formulas were detected in all four ionization modes, with approximately 24%, 15%, 6% and 2% being observed exclusively in +ESI, -ESI, +APPI and -APPI, respectively. Thus, the majority of molecular formulas in this study were detected in 2 or 3 of the ionization modes. The distribution of molecular formulas observed exclusively in one ionization mode are consistent with the expectations for these methods (Figs. 4.6 and 4.7) (Cech and Enke 2001, Kauppila et al. 2002, Tfaily et al. 2015). Formulas exclusive to -ESI were shifted to higher O/C (more oxidized), those exclusive to +ESI were shifted to lower O/C (less oxidized). Formulas exclusive to -APPI had moderate O/C and were shifted to lower H/C (more unsaturated) and those exclusive to +APPI were shifted to lower O/C and lower H/C (less oxidized and more unsaturated). Compared to previous studies of aerosol using only -ESI (Brege et al. 2018, Schum et al. 2018) the combination of these four ionization methods identifies 3 to 5 times the numbers of molecular formulas. This enhanced examination allows for the observation of the continuum of carbonaceous species in the complex mixture of organic aerosol.

The carbon continuum extends from aliphatic into the aromatic and condensed aromatic chemical space and we observe molecular formulas with proportionately high DBE over the entirety of observed carbon numbers. Many of the molecular formulas have AI_{mod} values defined for olefinic species while a notable number had AI_{mod} values suggesting an aromatic or condensed aromatic structure (Fig. 4.11). The average O/C ratio for the molecular formulas of BB05 was 0.363 ± 0.23 and slightly higher in BB06 with 0.368 ± 0.23 . Based on the limits described in Tu et al. (2016) an overwhelming majority of the molecular formulas have low to moderate oxygenation and oxidation, as shown by the distributions of formulas in Figs. 4.6 and 4.7 both left of the vertical $O/C = 0.6$ line and above the diagonal $OS_C = 0$ line. Although many of the molecular formulas were common to both samples, those formulas exclusively observed in BB06 compared to those exclusively observed in BB05 had slightly higher H/C (1.422 ± 0.41 vs. 1.217 ± 0.35) and O/C (0.352 ± 0.23 vs. 0.316 ± 0.26), indicating a slight enhancement in the atmospheric aging of BB06 (Fig. 4.12). These differences in oxidation are consistent with the shift towards larger particle sizes observed for BB06 (Fig. 4.4). The low oxygenation and high aromatic character observed here suggest the samples had an overall minimal influence of atmospheric aging, possibly explained by the high viscosity expected for tar balls and thus slow diffusion of atmospheric oxidants into the particles (Hinks et al. 2016, Reid et al. 2018, Schum et al. 2018).

Figure 4.11. Percent composition of molecular formulas in BB05 and BB06 by modified aromaticity index (AI_{mod}). AI_{mod} groups are based on calculated AI_{mod} values as described by Koch and Dittmar (2006, 2016) and these groups are defined in Table 4.1.

4.4.2 Volatility of tar ball molecular species

The volatility of the molecular formulas estimated using the C_0 model by Li et al. (2016) ranged from approximately 7.0×10^{-18} to $2.6 \times 10^7 \mu\text{g m}^{-3}$ and most molecular formulas in the samples were defined as low-volatile organic compounds (LVOC) and extremely low-volatile organic compounds (ELVOC) (Fig. 4.13). These estimates are in good agreement with the overall measured volatility of the aerosol particles. Fig. 4.14 shows the measured room-temperature evaporation kinetics of size-selected aerosol particles (150 nm, 234 nm and 311 nm) for BB05. Overall, the particles lost only $\sim 10\%$ of their volume after 24 h of evaporation in an organic-free environment, indicating that 90% of the particle volume was comprised of nearly non-volatile organics ($C_0 \leq 10^{-4} \mu\text{g m}^{-3}$). The calculated $T_{g,RH}$ values ranged from 155.6 to 423.3 K with the majority of formulas in both samples being distributed towards higher $T_{g,RH}$. There were noticeable differences in the calculated $T_{g,RH}$ values for molecular formulas observed in APPI compared to ESI, where the APPI formulas were shifted towards lower $T_{g,RH}$. This is especially noticeable in the formulas exclusive to an ionization method (upper panel of Figs 4.15 and 4.16) and especially for +APPI exclusive formulas. Molecular species expected to be more liquid-like would also be expected to have higher vapor pressure than solid species, which could play a role in their apparent enhanced detection with APPI methods here.

Based on the calculated ratios of $T_{g,RH}$ to T_{amb} , a majority of CHO formulas observed in BB05 are anticipated to be more highly viscous than those observed in BB06 (Fig. 4.17). These results suggest that the tar balls observed in the BB05 sample were solid spheres while those in BB06 were more likely to be semi-solid spheres, which is consistent with the trends in atmospheric aging described above by O/C and H/C ratios of sample exclusive molecular formulas. Because of the role of water as a plasticizer, this divergence in predicted phase may be due to the marginal differences in the ambient relative humidity during sample collection (see Table 4.2). These differences may also indicate the limits of the model used to calculate $T_{g,RH}$, which currently does not account for the significant influence CHNO species observed in these samples. The relationships between the estimated volatility, calculated $T_{g,RH}$ and molecular weight for BB05 and BB06 are shown in Figs 4.15 and 4.16 respectively. The carbon continuum appears to follow the expected trend of decreasing volatility with increasing $T_{g,RH}$ for increasing molecular weights. Formulas with higher $T_{g,RH}$ are anticipated to be in a semi-solid or solid state, as their $T_{g,RH}$ approach and exceed T_{amb} . The majority of the observed CHO molecular formulas were distributed towards lower volatility and higher $T_{g,RH}$, which is consistent with the expectation of low volatility and glassy-like phase states for tar balls (Posfai et al. 2004).

Figure 4.12. The distribution of all 11,841 molecular formulas observed of both BB05 and BB06 in van Krevelen space (center) with the density distribution of the formulas by O/C and H/C (top and right respectively) in order to visualize the magnitude of overlapping points in the center panel. Formulas exclusive to one sample or another are denoted with colored points or bars (BB05: blue and BB06: red) and formulas common to both samples appear in gray. Dashed lines represent $H/C = 1.2$ (vertical), $O/C = 0.6$ (horizontal) and $OS_C = 0$ (diagonal) (Tu et al. 2016).

Figure 4.13. Percent composition of molecular formulas in BB05 and BB06 by the estimated volatility groups. Volatilities are estimated via the calculated saturation mass concentration (C_0) as described by Li et al. (2016) and the volatility groups are defined in Table 4.1.

Figure 4.14. The measured room-temperature evaporation kinetics of size-selected aerosol particles (150 nm, 234 nm and 311 nm) for BB05. The particles lost ~10% of their volume after 24 h (1,440 min) of evaporation in an organic-free environment, indicating that ~90% of the particle volume was comprised of nearly non-volatile organics.

4.4.3 Health and atmospheric impacts

The catalytic generation of reactive oxygen species such as peroxides, superoxide anion and hydroxyl radical, have been associated with biomass burning aerosol and humic-like substances (Verma et al. 2015) and species with quinoid moieties (Lin and Yu 2011). Reactive oxygen species can cause oxidative stress in living organisms, making them, and by association biomass burning emissions, a sincere health hazard. We observe a significant number of aromatic and condensed aromatic species with low to moderate oxidation in our samples (Fig. 4.18) where quinones and species with quinoid moieties are possible. These types of aromatic and nitro-aromatic species could also be light absorbing and contribute to atmospheric brown carbon (Yu et al. 2016, Lin et al. 2017, Hawkins et al. 2018, Li et al. 2019).

Metabolic processing and oxidation of biomass burning emissions, such as PAH species, could catalyze the production of reactive oxygen species (Xue and Warshawsky 2005) and we expect many of the condensed aromatic species observed here could be related to PAHs. Lin et al. (2018) report the observation of molecular formulas in biomass burning samples which are likely to be PAHs, using a mix of extremely low-polarity extraction solvents. We do not observe any CH group molecular formulas with high enough AI_{mod} to indicate hydrocarbon PAHs, likely due to the ACN extraction used in this study. However, we do observe many CHO, CHN and CHNO species which are possibly functionalized oxo-PAHs and nitro-PAHs, with ring embedded heteroatoms also being possible. Functionalized PAH derivatives have been previously observed in aerosol, in correlation with PAHs, attributed to biomass burning sources (Samburova et al. 2016, Alves et al. 2017, Lin et al. 2018). While the specific carcinogenic effects of oxygen and nitrogen functionalized PAHs are still somewhat uncertain, in general terms, the negative health effects of absorbed PAHs occur once they have been oxidized within the biological system, allowing for polar interactions with cellular macromolecular structures like DNA and proteins (Xue and Warshawsky 2005). Functionalized PAHs would have increased polar character and water solubility, and thus the potential to be more biologically available for hazardous interactions than their hydrocarbon PAH counterparts.

A noticeable number of CHNO aromatic species observed here had high O/N ratios, making nitro-phenols and lignan-like oligomers possible. Many functionalized nitro-benzene compounds have known mutagenic effects (Shimizu and Yano 1986) and it is likely that exposure to biomass burning emissions like tar balls include these compounds.

Considering the observed carbon continuum of these samples, it is likely that exposure to the aromatic and condensed aromatic species described here for tar balls poses a significant health risk (Lin and Yu 2011, Verma et al. 2015, Alves et al. 2017). Wildfires are expected to increase in severity in North America with rising global temperatures (Westerling et al. 2006, Barbero et al. 2015, Gergel et al. 2017) and thus elevated emissions of tar balls may contribute to increased aerosol toxicity in addition to atmospheric warming from the production of absorbing aerosol.

Figure 4.15. Distribution of CHO molecular formulas for BB05 in estimated saturation mass concentration (C_0) (Li et al. 2016) vs. estimated relative humidity dependent glass transition temperature ($T_{g,RH}$) (DeRieux et al. 2018) space (bottom panel). The distribution of those molecular formulas by $T_{g,RH}$ (top panel) are color coded by the method of detection, whether a formula was exclusive to one of the four ionization modes (red, purple, blue, gold) detected in all four modes (green) or some other combination of multiple ionization modes (gray). The solid red line at 300 K represents the average ambient temperature (T_{amb}) during sample collection. Species with a $T_{g,RH}$ which approach or exceed T_{amb} are more likely to be in a semi-solid or solid phase respectively. (Shiraiwa et al. 2017, DeRieux et al. 2018). The dashed red line at 240 K represents the theoretical temperature for the expected phase change between liquid and semi-solid. The estimated volatility groups along the left side of the bottom panel are defined in Table 4.1.

Figure 4.16. Distribution of CHO molecular formulas for BB06 in estimated saturation mass concentration (C_0) (Li et al. 2016) vs. estimated relative humidity dependent glass transition temperature ($T_{g,RH}$) (DeRieux et al. 2018) space (bottom panel). The distribution of those molecular formulas by $T_{g,RH}$ (top panel) are color coded by the method of detection, whether a formula was exclusive to one of the four ionization modes (red, purple, blue, gold) detected in all four modes (green) or some other combination of multiple ionization modes (gray). The solid red line at 298 K represents the average ambient temperature (T_{amb}) during sample collection. Species with a $T_{g,RH}$ which approach or exceed T_{amb} are more likely to be in a semi-solid or solid phase respectively. (Shiraiwa et al. 2017, DeRieux et al. 2018). The dashed red line at 238 K represents the theoretical temperature for the expected phase change between liquid and semi-solid. The estimated volatility groups along the left side of the bottom panel are defined in Table 4.1.

Figure 4.17. Distribution of the ratio of the relative humidity glass transition temperature ($T_{g,RH}$) to the average ambient temperature during sample collection (T_{amb}), by the number of oxygen atoms in the CHO molecular formulas of BB05 and BB06. The ratio of $T_{g,RH}$ to T_{amb} can be used to estimate the phase state for a chemical species of a given molecular formula for established ambient conditions (Shiraiwa et al. 2017, DeRieux et al. 2018). The expected transition areas of 0.8 and 1.0 for this ratio are indicated in the Figure with dotted and dashed lines respectively (see Table 4.1).

Figure 4.18. Density distribution of oxygen to carbon ratios (O/C) of molecular formulas of BB05 (top) and BB06 (bottom). Formulas are separated by modified aromaticity index (AI_{mod}) (Koch and Dittmar 2006, 2016) and shape widths are scaled to the number of molecular formulas with the same O/C value. The horizontal line represents $O/C = 0.6$ (Tu et al. 2016).

The observed carbon continuum is consistent with the wide range of molecular formula compositions in several other atmospheric aerosol studies (Schmitt-Kopplin et al. 2010, Mazzoleni et al. 2012, Zhao et al. 2013, Kourtchev et al. 2014, Dzepina et al. 2015, Willoughby et al. 2016, Wang et al. 2017, Brege et al. 2018, Schum et al. 2018). It is likely that the observed carbon continuum is still incomplete and that the organic fraction of atmospheric aerosol is even more complex than described here. These results are influenced by both the ACN extraction technique for the aerosol filters, as well as the ionization methods used for analysis. Additional extraction methods using a polar solvent like water or a low polarity solvent like hexane could be performed to isolate different chemical species which were not extracted here by ACN. Additional ionization methods (such as atmospheric pressure chemical ionization) or the pre-separation of samples before analysis by liquid chromatography may also reveal additional layers of complexity to this carbon continuum. Considering that each molecular formula observed is likely composed of several different isomeric species (Zark et al. 2017, Hawkes et al. 2018) it is likely that the organic fraction of atmospheric aerosol is far more complex than any one analytical technique can reveal.

4.5 Associated content

4.5.1 Data availability

The final completed data set for this work is intended to be made available alongside the published results once the manuscript is completed and submitted.

4.5.2 Acknowledgements

Many supporting measurements and calculations for this chapter were prepared by our collaborators at the Pacific Northwest National Laboratory, Dr. Swarup China and Dr. Alla Zelenyuk, who have contributed ambient samples for analysis and have provided processed data. In the publication of this chapter they are intended to be listed as co-authors.

4.6 Supplementary information: chapter 4

4.6.1 Use of MFAssignR for molecular formula assignment

Mass spectra data was collected for two biomass burning samples (BB05 and BB06) using an ultrahigh resolution Orbitrap Elite mass spectrometer. These samples were significantly influenced by biomass burning tar ball particle morphology. Data was collected in the negative and positive polarities of electrospray ionization (ESI) and atmospheric pressure photoionization (APPI), hereafter abbreviated as -APPI, +APPI, -ESI, and +ESI. For full details of sample preparation and analysis, see section 4.3. Mass lists were generated from the collected mass spectra, using Xcalibur (Thermo Scientific v. 3.0). The spectra were collected for 200 scans, with the mass lists being the result of an integration over the areas

of steady ion signal. Data analysis of the mass lists was performed using R and RStudio (R v. 3.5.1 and RStudio v. 1.1.463), using a series of custom scripts. The recently developed open source MFAssignR (Schum et al. 2018) package for R was used to estimate the signal-to-noise ratio, recalibrate the mass spectra and assign molecular formulas to the measured masses. The full details of molecular formula assignment using MFAssignR are described elsewhere (Schum et al. 2018), but an in depth synopsis of the use of MFAssignR for these samples is provided in the following sections.

4.6.1.1 *Signal-to-noise estimation*

Noise levels were estimated using the “KMDNoise” function of MFAssignR. The function calculates Kendrick mass (KM) and Kendrick mass defect (KMD) for the entire input mass list. Ordinarily measured masses are relative to ^{12}C equaling exactly 12, however in the KM some other unit is set to an integer value. This other unit is commonly CH_2 which is set to equal exactly 14 (Stenson et al. 2003). The KM is calculated by multiplying the measured mass by the ratio of the integer mass of this new unit, to the exact monoisotopic mass of this new unit. An example of this calculation using CH_2 as the new base unit is provided in equation (4.1):

$$KM = \text{measured mass} * \frac{14.00000}{14.01565} \quad (4.1)$$

KMD is then calculated as the difference between the nominal mass (the sum of the integer masses of the elements in a molecular formula) and the calculated KM. Fig. 4.19 shows the KMD vs. measured ion mass output from the “KMDNoise” function for sample BB06 analyzed by -APPI. It can be seen that higher intensity analyte signal occupies a set space in the plot, and the surrounding area of lower intensity peaks can be used to estimate the noise. A diagonal slice of the noise area was used to calculate an average noise level for each sample, and a minimum peak intensity of six times this level was required for molecular formula assignment. Large gaps of signal are present in Fig. 4.19 (areas with no data points), especially above m/z 500, due to the spectrum averaging setting used during data collection, which reduces the overall noise of the collected spectra.

4.6.1.2 *Spectra recalibration*

The measured masses with intensities above the signal-to-noise threshold were recalibrated to lower the average absolute error of assigned molecular formulas and to assign additional formulas in the spectra. This process is described in greater detail elsewhere (Schum et al. 2018). Briefly, the recalibration estimates a mass correction for all ions in the spectra based on a chosen set of trusted masses (Kozhinov et al. 2013) and is further improved through segmented steps (Savory et al. 2011). A preliminary assignment of molecular formulas only allowing for carbon, hydrogen and oxygen (CHO) elements was used to generate a list of potential recalibrant series using the “RecalList” function. The “Racal” function was

Figure 4.19. Kendrick mass defect (KMD) vs. measured ion mass for sample BB06 analyzed by -APPI. This plot is part of the output of the “KMDNoise” function of MFAssignR and is used for signal-to-noise level determination. Point colors are scaled to the natural log of the peak intensity, in order to visually show the difference between higher intensity analyte peaks (yellow-red) and lower intensity noise peaks (teal-blue). The average intensity of the points between the two diagonal red bars was used to determine the signal-to-noise level, and a value of six times this level was required for a peak to be assigned a molecular formula.

then used with 4 to 6 CH₂ homologous seed series of high abundance. The recalibration process reduced the average absolute error of molecular formula assignment to below 1 ppm for all samples. The BB06 sample collected in +APPI did not undergo recalibration, because attempts to recalibrate made the average error for this sample higher. A list of the series used for recalibration in each sample, as well as the change in the average absolute error of molecular formula assignment, are provided in Table 4.5.

4.6.1.3 Molecular formula assignment

Following recalibration (with exception of BB06 analyzed by +APPI, which was not recalibrated) molecular formulas were assigned using the “MFAssign” function. As described above, only masses with intensities of six times the signal-to-noise ratio were considered for formula assignment. Formulas were assigned allowing for carbon, hydrogen, oxygen, nitrogen and sulfur, and up to one sodium atom for data collected in +ESI, allowing for positively charged sodium adducts in addition to hydrogen adducts. Limits were placed on oxygen (0 to 30), nitrogen (0 to 3) and sulfur (0 to 1) atoms, to reduce the number of poor formula assignments. Odd electron radicals were not considered for assignment. A de novo cut-off at m/z 300 was applied, indicating that no new formula assignments would occur above m/z 300, unless the formula was part of an existing homologous series that began at a point lower than m/z 300. Molecular formulas with the same KMD can be identified to extend formula assignments to higher masses, as the number of possible formulas increases greatly as molecular weight increases. These CH₂ homologous series are families of formulas with identical heteroatom content and DBE, and thus their assigned molecular formulas would only differ by some number of CH₂ units (this however does not necessarily represent a relationship in chemical structure). The CH₂ homologous series can be observed as horizontal rows of points in Fig. 4.19. Neutral masses were calculated by correcting for the respective hydrogen loss in negative mode, or a hydrogen adduct in positive mode, by adding or subtracting the exact mass of a proton (1.007276 Da). In the case of +ESI, sodium adducts were accounted for by correcting for the exact mass of Na⁺ (22.989222 Da). Measured masses were assumed to have a charge of ± 1 , as we did not expect strong influence from multiply charged species. The resulting molecular formulas were classified into the following elemental groups: "CH", "CHO", "CHN", "CHNO" and "CHOS" based on the respective inclusion of carbon, hydrogen, oxygen, nitrogen and/or sulfur elements in the formula. Classes of formulas can also be identified by heteroatom content (e.g. O₂ or NO₅) and DBE.

Table 4.5. List of recalibration series used as seeds, by sample and ionization method

Sample and ionization method	Seed series used		Average absolute error of preliminary formula assignment	Average absolute error of recalibrated formula assignment
	Class	DBE		
BB05: -APPI	O ₂	5	1.78 ppm	0.93 ppm
	O ₅	2		
	O ₄	7		
	O ₁₁	15		
BB05: +APPI	O ₂	5	1.36 ppm	0.95 ppm
	O ₄	9		
	O ₆	10		
	O ₇	13		
	O ₉	19		
BB05: -ESI	O ₄	2	1.16 ppm	0.57 ppm
	O ₅	4		
	O ₈	9		
	O ₁₂	9		
	O ₁₃	14		
BB05: +ESI	O ₂	5	1.61 ppm	0.62 ppm
	O ₄ [*]	4		
	O ₅ [*]	7		
	O ₁₃ [*]	12		
BB06: -APPI	O ₂	2	2.17 ppm	0.89 ppm
	O ₄	2		
	O ₄	3		
	O ₆	5		
	O ₁₀	10		
	O ₁₁	16		
BB06: +APPI	NA	NA	NA	NA
BB06: -ESI	O ₄	2	1.48 ppm	0.58 ppm
	O ₆	3		
	O ₈	9		
	O ₁₄	9		
BB06: +ESI	O ₂	5	1.41 ppm	0.54 ppm
	O ₄ [*]	4		
	O ₈ [*]	4		
	O ₈ [*]	11		

*Denotes series in +ESI analyses which contain formulas with Na⁺ adducts.

Figure 4.20. The distribution of molecular formulas by elemental class for BB06 analyzed by -APPI, separated into the CHO group (panel A) and the CHNO group (panel B). Please note the difference in scale of the y-axis between the two panels.

4.6.1.4 Data quality assurance and molecular formula validation

The absolute error (AE) of all assigned molecular formulas is defined in equation (4.2) and was limited to a value less than or equal to 3 ppm.

$$AE \text{ (ppm)} = \left| \frac{\text{measured mass} - \text{theoretical mass}}{\text{theoretical mass}} \right| * 10^6 \quad (4.2)$$

Only positive integer values, including 0, were allowed for DBE. Based on the work of Herzsprung et al. (2014) molecular formulas were also limited to a range for DBE minus oxygen (DBE-O) values of -13 to 13. DBE-O values were somewhat normally distributed over a continuum from -10 to 10; on an elemental group basis, clusters of formulas diverging from these normal distribution trends were removed. Elemental ratios for O/C and H/C were also used to validate the assigned formulas, only allowing for O/C ranging from 0 to 2.5 and for H/C ranging from 0.3 to 3. Formulas were validated with a “large atom check” also known as the rule of 13. A base hydrocarbon formula (C_nH_{n+r}) can be generated for any measured mass by solving: $\frac{\text{mass}}{13} = n + \frac{r}{13}$ (Pavia 2009). The maximum number of “large atoms” (C, O, N, S) in a formula is defined as the integer quotient of the mass divided by 13, as substituting for a heteroatom (O, N or S) in a molecular formula involves the substitution for at least one carbon atom. Thus, formulas exceeding this maximum number of “large atoms” can be rejected. In a similar manner, the maximum number of hydrogens can be determined for a molecular formula as $n+r$, and formulas exceeding this number can also be rejected. The nitrogen rule was used to remove formulas with odd masses that do not contain an odd number of nitrogen atoms, and even masses that do not contain an even number (or no) nitrogen atoms, due to the odd numbered valence of nitrogen (Pavia 2009).

Table 4.6. Excerpt of ambiguous molecular formula assignments for BB06 analyzed by -APPI.

Measured mass (Da)	Molecular formula	Elemental class	Group	Theoretical mass (Da)	Absolute Error (ppm)
344.0926	C ₂₁ H ₁₅ NO ₄ *	NO ₄	CHNO	344.0928	0.696
344.0926	C ₁₃ H ₁₉ N ₃ O ₆ S	N ₃ O ₆ S	CHNOS	344.0922	1.193
372.1234	C ₂₃ H ₁₉ NO ₄ *	NO ₄	CHNO	372.1241	1.895
372.1234	C ₁₅ H ₂₃ N ₃ O ₆ S	N ₃ O ₆ S	CHNOS	372.1235	0.148
386.1031	C ₂₃ H ₁₇ NO ₅ *	NO ₅	CHNO	386.1034	0.869
386.1031	C ₁₅ H ₂₁ N ₃ O ₇ S	N ₃ O ₇ S	CHNOS	386.1027	0.815
400.1543	C ₂₅ H ₂₃ NO ₄ *	NO ₄	CHNO	400.1554	2.858
400.1543	C ₁₇ H ₂₇ N ₃ O ₆ S	N ₃ O ₆ S	CHNOS	400.1548	1.233
402.1345	C ₂₄ H ₂₁ NO ₅ *	NO ₅	CHNO	402.1347	0.584
402.1345	C ₁₆ H ₂₅ N ₃ O ₇ S	N ₃ O ₇ S	CHNOS	402.1340	1.033

*Denotes the molecular formulas with NO₄ and NO₅ elemental classes which were ultimately chosen to remain in the data set, opposed to the formulas with N₃O₆S and N₃O₇S elemental classes.

The data set was processed to remove formulas in CH₂ homologous series with large gaps, as well as CH₂ homologous series with a length of one. Because of the mechanics of ionization for -ESI, we do not expect O = 0 species (CH and CHN) to be detected by this method, and they were removed from the resulting data set. The distribution of formulas by elemental class followed a somewhat normal trend (Fig. 4.20), and formulas which notably diverged from these trends were also removed. For example, a formula of classes O₁₉ or N₃O₂ do not fit into the trends observed for BB06 in Fig. 4.20.

Ambiguous formula assignments (multiple molecular formulas assigned to the same measured mass) were resolved manually. Most ambiguous formula assignments can be internally validated by MFAssignR using homologous series trends, as described above, however some require manual review by the user. There were approximately 50 to 300 ambiguous formula assignments per analysis technique and sample, with usually 2 ambiguous assignments per mass. These ambiguous assignments had thus far passed all previous forms of formula validation (O/C and H/C ratios, DBE-O, rule of 13, nitrogen rule, etc.). Trends in the distribution of unambiguous assignments (measured masses with only one viable molecular formula assignment) were used as a means of validation for ambiguous assignments. The distribution of the unambiguous formulas by elemental class followed a somewhat normal trend (Fig. 4.20), so ambiguous formulas of classes that did not fit into these trends could be invalidated. We observed regular patterns for sets of ambiguous assignments across different measured masses, an example of which is provided in Table 4.6 for BB06 analyzed by -APPI. In Table 4.6, there is a trend of NO_z vs. N₃O_{z+2}S ambiguous formulas. There is a 0.6519 mDa difference, or “mass shift”, between the theoretical masses of these two formula classes, allowing for the substitution of C₈ for H₄N₂O₂S, while remaining within the 3 ppm absolute error tolerance of formula

assignment. In this example, we were able to validate the NO_z class of formulas based on the formula class distributions shown in Fig. 4.20 for CHNO species. In contrast, there were no other CHNOS formulas detected in this sample, so we were suspicious of these CHNOS formulas with N₃, especially without the observation of CHNOS formulas with N₁ and N₂. It can be seen in Table 4.6 that choosing the ambiguous formula with the lowest absolute error was not necessarily the best choice. Other ambiguous formula assignments were validated in a similar manner across the other ionization methods, as well as for BB05.

4.6.1.5 Blank subtraction

Filter extraction, sample preparation, mass spectrometry analysis and molecular formula assignment was also performed on a blank filter sample. Formula assignment was carried out for the blank sample, in the same manner as described above for BB05 and BB06. On an individual basis, molecular formulas in the samples were compared to the respective blank and matching molecular formulas with a higher abundance in the blank compared to the sample were removed from the data set.

4.6.2 Estimation of chemical and physical properties

A series of calculations were used to estimate chemical and physical properties for each molecular formula, as described in the following sub sections.

4.6.2.1 Double bond equivalents

The number of double bonds and ring structures in a molecular formula, or double bond equivalents (DBE), was estimated by equation (4.3) using the molecular formula format C_cH_hO_oN_nS_s.

$$DBE = c - \left(\frac{h}{2}\right) + \left(\frac{n}{2}\right) + 1 \quad (4.3)$$

4.6.2.2 Aromaticity index

The modified aromaticity index (AI_{mod}) described by Koch and Dittmar (2006, 2016) was used to estimate the aromatic character of the formulas, and was calculated using equations (4.4) to (4.6) using the same molecular formula format as DBE.

$$DBE_{AI,mod} = 1 + c - \frac{o}{2} - s - \left(\frac{n+h}{2}\right) \quad (4.4)$$

$$C_{AI,mod} = c - \frac{o}{2} - n - s \quad (4.5)$$

$$AI_{mod} = \frac{DBE_{AI,mod}}{C_{AI,mod}} \quad (4.6)$$

In equation (4.6), the $AI_{\text{mod}} = 0$, if $DBE_{AI,\text{mod}} \leq 0$ or $C_{AI,\text{mod}} \leq 0$, as defined by Koch and Dittmar (2006, 2016).

AI_{mod} was calculated as shown in equations (4.4) to (4.6) for all CH, CHN and CHO formulas as well as positive mode CHNO formulas, assuming that half of the oxygen atoms in the formula are bound in carbonyls. For the AI_{mod} calculation of negative mode CHNO and CHOS formulas, equations (4.4) and (4.5) were adjusted to account for a number of oxygen atoms in the formula which are assumed to be bound in nitro, nitrate, sulfonate and sulfate groups, by substituting the $\left(\frac{o}{2}\right)$ term for $\left(\frac{o}{2} - 3\right)$. Hereafter, and in the main text, both calculations are collectively referred to as AI_{mod} . The defined ranges for AI_{mod} values are described in Table 4.1.

4.6.2.3 Average oxidation state of carbon

The average oxidation state of carbon (OS_C) was calculated based on the approximation described by Kroll et al. (2011) using equation (4.7). Equation (4.8) is a modified version of equation (4.7) and was used for negative mode CHNO and CHOS formulas, to account for the electrons bound in expected nitro, nitrate, sulfonate and sulfate groups.

$$OS_C \approx 2\frac{o}{c} - \frac{h}{c} \quad (4.7)$$

$$OS_{C_{\text{mod}}} \approx 2\frac{o}{c} - \frac{h}{c} - 5\frac{n}{c} - 6\frac{s}{c} \quad (4.8)$$

Equation (4.8) assumes that nitrogen and sulfur atoms in the formula are fully oxidized, which is a reasonable assumption for negative mode data. As with AI_{mod} , these different calculations are collectively referred to as OS_C from here on, and in the main text.

4.6.2.4 Volatility estimation

The volatility of the assigned molecular formulas was estimated based on the work of Li et al. (2016) The saturation mass concentration (C_0) (units of $\mu\text{g m}^{-3}$) was calculated using equation (4.9) as described by (Li et al. 2016)

$$\text{Log}_{10}C_0 = (n_C^0 - n_C)b_C - n_O b_O - n_H b_H - 2\frac{n_C n_O}{n_C + n_O} b_{CO} - n_N b_N - n_S b_S \quad (4.9)$$

In equation (4.9), n_C , n_O , n_H , n_N and n_S are the respective number of carbon, oxygen, hydrogen, nitrogen and sulfur atoms of a given molecular formula, n_C^0 is the reference carbon number, b_{CO} is the carbon-oxygen nonideality parameter, and b_C , b_O , b_H , b_N and b_S are the respective empirical best fit parameters for the contributions of carbon, oxygen, hydrogen, nitrogen and sulfur atoms to the calculation of $\text{Log}_{10}C_0$. The parameters used for

Table 4.7. Estimated volatility calculation parameters used in equation (4.9) as described in Li et al. (2016).

Elemental Group	n_C^0	b_C	b_H	b_O	b_{CO}	b_N	b_S
CH	17.95	0.5742	-0.1417	0	0	0	0
CHO	15.77	0.6238	-0.1387	1.735	-0.8592	0	0
CHN	23.01	0.4307	-0.0211	0	0	0.9528	0
CHNO	21.12	0.4139	-0.0376	0.8092	-0.1174	1.1010	0
CHOS	16.07	0.5348	-0.1507	1.354	-0.4175	0	0.8993

equation (4.9) are provided in Table 4.7, as described in Li et al. (2016). These parameters were obtained using multi-linear least squares method for the best fit of a set of 30,000 compounds (CH, CHO, CHN, CHNO, CHOS and CHNOS species), at a temperature of 298 K (Li et al. 2016). This temperature is a reasonable approximation of the average temperatures during sample collection for BB05 and BB06 (Table 4.2), so the temperature correction described in Li et al. (2016) was not utilized. The defined ranges for C_0 values are described in Table 4.1.

4.6.2.5 Dry and relative humidity dependent glass transition temperatures

The dry glass transition temperature ($T_{g,dry}$) was estimated for CHO formulas using equation (4.10) as described by DeRieux et al. (2018).

$$T_{g,dry} = \left(n_C^0 + \ln(n_C) \right) b_C + \ln(n_H) b_H + \ln(n_C) \ln(n_H) b_{CH} + \ln(n_O) b_O + \ln(n_C) \ln(n_O) b_{CO} \quad (4.10)$$

In equation (4.10), n_C , n_O and n_H are the respective number of carbon, oxygen and hydrogen of a molecular formula, n_C^0 is the reference carbon number, b_C , b_O and b_H , are the respective empirical best fit parameters for the contributions of carbon, oxygen and hydrogen atoms respectively to the calculation of $T_{g,dry}$, and b_{CH} and b_{CO} represent the contributions of carbon-hydrogen and carbon-oxygen bonds respectively. The parameters used for equation (4.10) are provided in Table 4.8, as described in DeRieux et al. (2018)

We then used the Gordon-Taylor equation to estimate the relative humidity dependent glass transition temperature ($T_{g,RH}$) of organic-water mixtures (Shiraiwa et al. 2017, DeRieux et al. 2018) as described in equation (4.11).

$$T_{g,RH} = \frac{(1-w_{org})T_{g,w} + \frac{1}{k_{GT}} * w_{org} * T_{g,dry}}{(1-w_{org}) + \frac{1}{k_{GT}} * w_{org}} \quad (4.11)$$

Table 4.8. Estimated dry glass transition temperature ($T_{g,dry}$) calculation parameters used in equation (4.10) as described in DeRieux et al. (2018).

Elemental Group	n_c^0	b_c	b_H	b_{CH}	b_o	b_{co}
CHO	12.13	10.95	-41.82	21.61	118.96	-24.38

In equation (4.11), w_{org} is the mass fraction of organics, $T_{g,w}$ is the glass transition temperature for water (136 K), k_{GT} is the Gordon-Taylor constant (equal to 2.5) (Shiraiwa et al. 2017, DeRieux et al. 2018) and $T_{g,dry}$ is the estimated dry glass transition temperature calculated by equation (4.10). As described in Shiraiwa et al. (2017), DeRieux et al. (2018) and our previous work Schum et al. (2018) we estimated w_{org} by equation (4.12).

$$w_{org} = \frac{1.4 - \frac{1.4 \cdot RH}{100}}{1.4 - \frac{1.28 \cdot RH}{100}} \quad (4.12)$$

A ratio of $T_{g,RH}$ to the ambient temperature during sample collection (T_{amb}), was used to estimate the phase state ratio (PSR) of the molecular formulas as described in equation (4.13).

$$PSR = \frac{T_{g,RH}}{T_{amb}} \quad (4.13)$$

The defined ranges for PSR values are described in Table 4.1.

4.6.3 Molecular formulas observed in multiple ionization methods

Approximately 4,000 to 7,000 molecular formulas were assigned to each sample for each of the four ionization modes. Many molecular formulas were detected in multiple ionization modes, and we emphasize that the same molecular formula detected across different ionization methods may very likely represent different structurally distinct chemical species due to the different mechanisms for ionization involved. We define 15 orthogonal fractions of the total number of individual molecular formulas observed in a sample:

- 1) Exclusive to -APPI
- 2) Exclusive to +APPI
- 3) Exclusive to -ESI
- 4) Exclusive to +ESI
- 5) Universally common to all ionization methods
- 6) Common to both negative ionization modes
- 7) Common to both positive ionization modes
- 8) Common to both APPI analyses
- 9) Common to both ESI analyses
- 10) Intersecting -ESI, +ESI and -APPI ($\cap_1$)
- 11) Intersecting -ESI, +ESI and +APPI ($\cap_2$)

- 12) Intersecting -APPI, +APPI and -ESI ($\cap_3$)
- 13) Intersecting -APPI, +APPI and +ESI ($\cap_4$)
- 14) Intersecting -ESI and +APPI ($\cap_5$)
- 15) Intersecting +ESI and -APPI ($\cap_6$)

The six intersecting fractions (in the list above as numbers 10 to 15) are abbreviated as $\cap_1$ through $\cap_6$ respectively. These fractions are described graphically with a Venn diagram in Fig. 4.21 for BB05 and Fig. 4.22 for BB06. There was good agreement between BB05 and BB06 as to the distribution of the observed formulas across these different fractions. Overall, approximately 16% of formulas were detected in all four ionization methods, with 15% exclusive to -ESI, 24% exclusive to +ESI, 2% exclusive to -APPI and 6% exclusive to +APPI. Following the validation procedures described in section 4.6.1 and accounting for the overlap between ionization methods described above, there were 10,887 molecular formulas detected in BB05 and 10,553 molecular formulas detected in BB06. Further accounting for overlapping formulas between the two samples, 11,841 molecular formulas were identified for biomass burning tar ball influenced atmospheric aerosol.

BB05 total number of formulas: 10887

Figure 4.21. Venn diagram of the orthogonal fractions of molecular formulas observed in BB05. Circle areas are not to scale. The total number of formulas observed by each ionization method are provided adjacent to the respective method (outside the circle area) while the number of formulas associated with individual fractions are provided inside the fraction area next to the name (with the exception for the intersecting fractions, $\cap_1$ to $\cap_6$, which are provided in the area to the left). These intersecting fractions are further defined in section 4.6.3.

BB06 total number of formulas: 10553

Figure 4.22. Venn diagram of the orthogonal fractions of molecular formulas observed in BB06. Circle areas are not to scale. The total number of formulas observed by each ionization method are provided adjacent to the respective method (outside the circle area) while the number of formulas associated with individual fractions are provided inside the fraction area next to the name (with the exception for the intersecting fractions, n_1 to n_6 , which are provided in the area to the left). These intersecting fractions are further defined in section 4.6.3.

5 Conclusion

5.1 Summary

In this work we explored the molecular formula composition of atmospheric aerosol and fog. The presence of organic aerosol in the atmosphere has important implications for both the global climate as well as human health. Overall, the complex mixture of organic compounds is poorly characterized due to the challenges associated in resolving the molecular components for analysis. Ultrahigh resolution mass spectrometry was used in this work and we have shown an extreme complexity of molecular species to be present.

The analysis of biomass burning influenced aerosol and fog water from Po Valley revealed molecular trends for chemical transformations occurring in fog and aerosol liquid water. The molecular formulas extracted from aerosol were distributed towards higher carbon numbers compared to those of fog. Conversely, the molecular formulas from fog were distributed towards higher oxygen numbers than aerosol. These trends are consistent with the aqueous phase processes expected for fog water which can lead to enhanced fragmentation and oxidation (Ervens et al. 2011, Kroll et al. 2011). Functionalized species were observed in Po Valley fog and aerosol with N and S-containing formulas, expected to be oxidized organonitrates, organosulfates and nitrooxy-organosulfates. The high number of S-containing formulas and the similarity between the “fresh” influenced fog water and the “aged” influenced aerosol samples were somewhat unexpected results. These S-containing species were expected to have been formed in aerosol liquid water and degraded in fog with high liquid water content and low pH (Darer et al. 2011, Hu et al. 2011).

Our expanded work on Po Valley ambient fog and aerosol samples included multiple ionization methods and a much larger sample set. Here we observed many of the same trends from our previous study regarding the presence of N and S-containing formulas and carbon number distribution trends. The enhanced intensity of CHN molecular formulas in nighttime aerosol samples was an unexpected trend. We also observed an enhanced presence of N-containing molecular species in terms of both the number of molecular formulas and the detected ion intensities, possibly due to the inclusion of positive mode electrospray ionization data. We expect many of these N-containing molecular formulas to be aromatic and could thus contribute to atmospheric brown carbon. Furthermore, we have compared the ambient data to a laboratory generated sample expected to contain pyrazine-based chromophores (Hawkins et al. 2018), where we see a strong overlap in the molecular formulas of the two data sets.

The analysis of wildfire aerosol from the Pacific Northwest, expected to contain atmospheric tar balls, has been our most comprehensive study to date. Here we employed a suite of different soft-ionization methods to observe the most molecular level detail possible, as tar ball chemistry is poorly characterized. We observed that each additional ionization method provided some level of detail to the overall molecular characterization

which was not observed by the other methods. The resulting data set was vastly complex and showed a continuum of carbonaceous species in organic aerosol. The estimated volatilities and humidity dependent glass transition temperatures (Li et al. 2016, DeRieux et al. 2018) calculated for the observed molecular formulas was overall consistent with the low volatility and solid phase state expected for tar ball aerosol particles.

5.2 Recommendations for further research

Overall, in each study there was a significant fraction of molecular formulas which were common between the respective samples. These similarities may be an inherent aspect of the organic mixture in atmospheric aerosol. In this work the extreme differences between samples were focused on by analyzing and discussing the molecular formulas which were exclusive to one sample or another or the difference intensity trends between samples which highlight these distinctions. It may be an interesting avenue of research to instead focus on those common molecular formulas between samples and to determine if they are indeed related chemical compounds and if so their origins to the atmosphere.

Our direct observations thus far have shown that approximately 2,000 to 5,000 molecular formulas can be observed per sample, per ionization method used, with extraction method also playing some role in the molecular species that are observed. This work currently provides no estimation of the number of structural isomers associated with each molecular formula. Literature sources estimate for similar systems of complex organic matter that the number of chemical species present is at least one order of magnitude larger than the number of molecular formulas observed (Zark et al. 2017, Hawkes et al. 2018). The question then remains, how much molecular detail can be observed for organic aerosol with mass spectrometry?

The method employed in chapter 3 with overlapping scan ranges provided ~5 to 10% additional molecular formulas per sample, especially at higher molecular weights. This process requires a minor amount of effort in data acquisition and analysis and should be continued in future studies using the Orbitrap Elite instrument. The stitching of mass spectra over multiple different scan ranges may also be possible (Gavard et al. 2019), though there would be further reduced confidence in the intensity of the detected ions and increased complexity in data analysis.

The use of complementary overlapping extraction methods like the water extraction coupled to solid phase extraction employed in chapters 2 and 3, and the simple solvent phase extraction employed in chapter 4, would be a good way to observe orthogonal fractions of organics in ambient samples. While these different extraction methods are somewhat low effort to prepare, each would provide additional compounding difficulty to data acquisition and analysis.

Tandem mass spectrometry (MS^2) methods could be employed to analyze the molecular fragments of analyte molecules. While ultrahigh resolution methods provide reliable accurate mass measurements for molecular formula assignment, they provide no chemical

structural information or functional group characteristics. The inclusion of MS² methods would require some additional effort in data acquisition and a significant challenge in data analysis. However, these efforts would be worth the investment for reliable structural and functional group information to complement molecular formula assignments. The analysis of MS² data could eventually be automated in a similar fashion to molecular formula assignment now performed with custom R scripts.

The inclusion of additional ionization modes, as in chapters 3 and 4, is a non-trivial process. While these additional ionization modes are overall worth the effort, they should be employed situationally as each additional ionization mode compounds the amount effort required for both data acquisition and analysis substantially. It may be possible to employ ionization modes not discussed here, such as atmospheric pressure chemical ionization (APCI). While APCI does generate some radical species in the ionization process and can cause some fragmentation, this ionization mode may be more consistent than electrospray ionization with changing solvent compositions and thus has potential for the application of liquid chromatographic (LC) analysis methods.

The coupling of ultrahigh resolution mass measurements to LC separation methods is an ambitious but worthwhile goal. The development of LC methods would require a significant time investment and would also increase the burden of data analysis substantially. However, the advantages of a functioning LC analysis methods outweigh these costs. The pre-separation of analytes by LC would help to diminish charge competition using electrospray during data acquisition, as efficiently ionizing species could be isolated from less efficiently ionizing species. This would lead to less ion suppression by charge competition and the ability to observe a greater fraction of organics in samples. LC methods also have the potential to resolve isomer species, even if only to note that multiple versions of the same molecular formulas were detected. An example of isomeric molecular formula separation using LC is shown in Figs. 6.12 and 6.13. Simultaneous absorption measurements (such as UV-vis or diode array detectors) could be used to relate mass spectra to absorption and possibly establish quantitation. Separated ions analyzed by MS² would also be easier to interpret from an initial reduced complexity of ions to be fragmented. Overall, LC analysis methods are a respectable goal for the analysis of ambient aerosol, but there are currently many difficult challenges to establish these methods. Due to significant dilution of the sample during analysis ambient samples would also have to be fairly concentrated for LC methods, which may or may not be possible with some samples.

A hybrid method somewhere between LC and direct infusion may be the next step in aerosol analysis by mass spectrometry. A hybrid method could use a short chromatography column and high organic solvent composition, being less focused on fully resolving the organic matter and more focused on a general separation of analytes (polar from non-polar and ionic salts from organics). The automation of data acquisition using an LC autosampler would also be invaluable, even if a method does not use a chromatography column and essentially performs an automated direct infusion experiment. Automation would reduce the time investment for the analysis of many samples. A researcher's time could then be

focused on data analysis. What is learned from these hybrid methods could be invaluable for establishing a full LC method.

Nearly all the modifications to analysis techniques described here would increase the molecular level of detail observed with mass spectrometry analyses but would also increase the burden of data analysis. The combination of multiple modifications (such as simultaneous additional ionization methods and additional extraction techniques) compound these burdens multiplicatively. Though the formula assignment process and post-processing have been greatly streamlined using R scripts and MFAssignR, the increased size of the resulting data sets is a concern. It may be useful to develop advanced statistical analyses for these resulting large data sets to aid in data interpretation. With sufficient sample variables and number of samples, principle component analysis may reveal additional details about assigned molecular formulas. It may also be useful to develop machine learning models to aid in data analysis workflows. Automated algorithms to manage large data sets may be useful in determining signal to noise thresholds, choosing recalibration seed series, the validation of molecular formulas and resolving ambiguous assignments. Especially for the analysis of MS² and LC-MS data sets, the need for data reduction methods is critical. Developing these algorithms correctly may require a significant time investment however, as there exists a sincere difficulty barrier between computer science methods and traditional data analysis for analytical chemistry.

Overall, there are many options for improved molecular level data analysis of ambient aerosol and currently no one analytical method can resolve the complete mixture. In this work we have established that multiple analysis methods, as well as sample preparation methods, reveal additional layers of complexity. Even with the methods described here, the complex mixture in organic aerosol is far from being fully resolved, and far from being fully understood.

6 Appendix A: ancillary tables and figures

6.1 Ancillary tables

Table 6.1. Molecular formulas observed in chapter 2 Po Valley samples which were also observed in biomass burning influenced ambient cloud water samples from Mt. Tai, China as described in Desyaterik et al. (2013).

Formula ^a	SPC0106F ^b	SPC0201F ^b	BO0204N ^b	BO0213D ^b	Possible Identity
C ₅ H ₆ O ₃	ND	ND	ND	Low	
C ₅ H ₆ O ₄	ND	Low	ND	Low	
C ₅ H ₈ O ₄	ND	Med	ND	Low	methyl-succinic acid and glutatric acid
C ₅ H ₈ O ₅	ND	ND	ND	Low	Hydroxy-glutaric acid
C ₆ H ₅ NO ₃	High	High	Med	Med	nitrophenol
C ₆ H ₅ NO ₄	High	High	Very High	Med	nitrocatechol
C ₆ H ₅ NO ₅	ND	ND	Med	Low	
C ₆ H ₈ O ₄	Low	Med	ND	Med	
C ₆ H ₁₀ O ₃	ND	Low	ND	Low	
C ₆ H ₁₀ O ₄	ND	Med	ND	Low	methyl-glutaric acid and adipic acid
C ₆ H ₁₀ O ₅	ND	Low	ND	Low	levo-glucosan
C ₇ H ₆ O ₂	Med	Med	ND	Low	benzoic acid
C ₇ H ₆ N ₂ O ₅	Med	Med	Med	Low	
C ₇ H ₆ N ₂ O ₆	Med	ND	Med	Low	
C ₇ H ₇ NO ₃	High	High	Med	High	methyl-nitrophenol
C ₇ H ₇ NO ₄	High	Med	Med	Med	nitro-guaiacol and methyl-nitrocatechol
C ₇ H ₇ NO ₅	Med	Low	Med	Med	
C ₈ H ₅ NO ₄	Med	Low	Med	Low	
C ₈ H ₆ O ₃	Med	Med	Low	Low	
C ₈ H ₆ O ₄	Med	Very High	Low	Low	phthalic acid
C ₈ H ₇ NO ₃	Med	ND	Med	High	

Formula^a	SPC0106F^b	SPC0201F^b	BO0204N^b	BO0213D^b	Possible Identity
C ₈ H ₇ NO ₄	Med	Med	Med	High	
C ₈ H ₇ NO ₅	High	Med	Med	High	
C ₈ H ₈ O ₂	Med	Med	Low	Low	o-toluic acid
C ₈ H ₈ O ₃	High	Med	Med	Low	vanillin
C ₈ H ₉ NO ₃	Med	Med	Med	Low	
C ₈ H ₉ NO ₄	High	Med	Very High	High	
C ₈ H ₉ NO ₅	High	Med	High	High	
C ₉ H ₇ NO ₄	Med	Low	Med	Low	
C ₉ H ₈ O ₂	Med	Low	Low	ND	
C ₉ H ₈ O ₃	Med	Med	Med	Med	coumaric acid
C ₉ H ₉ NO ₃	Med	ND	Low	Med	
C ₉ H ₉ NO ₄	Med	Low	Med	Med	
C ₉ H ₁₁ NO ₄	Med	Low	Med	Med	
C ₁₂ H ₁₀ N ₂ O ₈	ND	ND	Low	ND	

^aAs the identified molecular formulas likely represent a variety of structural isomers, we note that matched molecular formulas do not necessarily correspond to the same molecular structure or atmospheric origin. ^bThe normalized abundances are indicated for each sample, with: “ND” (not detected), “Low” ($\leq 3\%$), “Med”, ($> 3\%$ and $\leq 15\%$), “High” ($> 15\%$ and $\leq 50\%$) and “Very High” ($> 50\%$).

Table 6.2. Molecular formulas observed in chapter 2 Po Valley samples which were also observed in the products of laboratory phenolic aqueous SOA reactions as described in Yu et al. (2014) and Yu et al. (2016).

Formula^a	SPC0106F^b	SPC0201F^b	BO0204N^b	BO0213D^b	Possible Identity
C ₄ H ₆ O ₅	ND	ND	ND	Low	
C ₅ H ₆ O ₅	ND	ND	ND	Low	ketoglutaric acid
C ₆ H ₆ O ₅	ND	ND	ND	Low	
C ₆ H ₈ O ₆	ND	ND	ND	Low	
C ₇ H ₆ O ₃	Med	Med	Med	Med	dihydroxy-benz-aldehyde
C ₇ H ₆ O ₄	Med	Low	Low	Med	
C ₇ H ₆ O ₅	ND	Low	ND	Low	
C ₇ H ₁₀ O ₆	Low	Low	Low	Med	
C ₇ H ₁₂ O ₇	Low	ND	ND	ND	
C ₈ H ₆ O ₅	Med	Med	ND	Low	
C ₈ H ₁₀ O ₃	Med	Low	ND	Low	syringol
C ₈ H ₁₀ O ₅	Med	Med	Low	Med	
C ₈ H ₁₀ O ₆	Low	Med	Low	Med	
C ₈ H ₁₀ O ₇	ND	Low	ND	Low	
C ₉ H ₁₀ O ₃	Med	Low	Low	Low	aceto-vanillone
C ₉ H ₁₀ O ₄	Med	Med	Low	Low	syring-aldehyde
C ₁₀ H ₈ O ₃	Med	Low	Med	Low	
C ₁₁ H ₁₀ O ₈	ND	Low	ND	ND	
C ₁₂ H ₁₀ O ₂	Low	ND	Low	ND	
C ₁₂ H ₁₀ O ₃	Low	ND	Low	Low	
C ₁₂ H ₁₀ O ₄	ND	ND	Low	Low	
C ₁₂ H ₁₀ O ₇	Med	Med	Low	Low	
C ₁₂ H ₁₂ O ₆	Med	Med	Low	Low	
C ₁₂ H ₁₂ O ₇	Med	Med	Low	Low	
C ₁₂ H ₁₄ O ₄	Med	Low	Low	Med	
C ₁₃ H ₁₀ O ₃	ND	ND	ND	Low	
C ₁₃ H ₁₀ O ₄	Med	Low	Low	ND	
C ₁₃ H ₁₀ O ₅	Med	ND	Low	Low	
C ₁₃ H ₁₂ O ₄	ND	ND	Low	Low	
C ₁₃ H ₁₂ O ₆	Med	Med	Low	Low	
C ₁₃ H ₁₄ O ₅	Med	Med	Med	Med	

Formula^a	SPC0106F^b	SPC0201F^b	BO0204N^b	BO0213D^b	Possible Identity
C ₁₃ H ₁₄ O ₇	Med	Med	Low	Low	
C ₁₃ H ₁₆ O ₈	Med	ND	Low	ND	
C ₁₄ H ₁₀ O ₅	Med	Low	Low	Low	
C ₁₄ H ₁₂ O ₆	Med	Low	Low	Low	
C ₁₄ H ₁₂ O ₇	Med	Med	ND	Low	
C ₁₄ H ₁₄ O ₄	ND	ND	Med	Low	
C ₁₄ H ₁₄ O ₅	Med	Low	Low	Med	
C ₁₄ H ₁₄ O ₆	Med	Med	Low	Med	
C ₁₄ H ₁₄ O ₈	Med	Med	ND	Low	
C ₁₄ H ₁₆ O ₈	Med	Med	Low	Low	
C ₁₄ H ₁₆ O ₉	Med	Low	ND	Low	
C ₁₄ H ₁₆ O ₁₀	Low	Low	ND	Low	
C ₁₅ H ₁₄ O ₆	Med	Low	Low	Low	
C ₁₅ H ₁₄ O ₈	Med	Med	Low	Low	
C ₁₅ H ₁₆ O ₆	Med	Med	Low	Low	
C ₁₅ H ₁₆ O ₈	Med	Med	ND	Low	
C ₁₅ H ₁₆ O ₉	Med	Med	ND	Low	
C ₁₅ H ₁₈ O ₇	Med	Med	Low	Low	
C ₁₅ H ₁₈ O ₉	Med	Med	Low	Low	
C ₁₅ H ₁₈ O ₁₀	Low	Low	ND	Low	
C ₁₆ H ₁₈ O ₆	Med	Low	Low	ND	
C ₁₆ H ₁₈ O ₇	Med	Med	Low	Low	
C ₁₆ H ₁₈ O ₉	Med	Med	ND	Low	
C ₁₈ H ₁₂ O ₅	Low	ND	ND	Low	
C ₁₈ H ₁₄ O ₄	Low	ND	Low	Low	
C ₂₀ H ₁₄ O ₆	Low	ND	Low	ND	
C ₂₀ H ₁₆ O ₇	Med	ND	Low	Low	
C ₂₀ H ₁₈ O ₆	Med	ND	Low	Low	
C ₂₁ H ₁₈ O ₈	Med	Low	Low	Low	
C ₂₁ H ₂₀ O ₆	Low	ND	Low	Low	
C ₂₁ H ₂₀ O ₈	Med	ND	Low	ND	
C ₂₈ H ₂₆ O ₈	ND	ND	Low	Low	

^aAs the identified molecular formulas likely represent a variety of structural isomers, we note that matched molecular formulas do not necessarily correspond to the same molecular structure or atmospheric origin. ^bThe normalized abundances are indicated for each sample, with: “ND” (not detected), “Low” ($\leq 3\%$), “Med”, ($> 3\%$ and $\leq 15\%$), “High” ($> 15\%$ and $\leq 50\%$) and “Very High” ($> 50\%$).

Table 6.3. Molecular formulas observed in chapter 2 Po Valley samples with expected influence from terpene oxidation products, which were also observed in biomass burning influenced cloud water as described in Zhao et al. (2013) and Cook et al. (2017).

Formula^a	SPC0106F^b	SPC0201F^b	BO0204N^b	BO0213D^b	Possible Identity
C ₇ H ₁₀ O ₄	Low	Med	Low	Med	
C ₇ H ₁₂ O ₆ S	ND	Low	ND	Low	
C ₇ H ₁₄ O ₅ S	Low	Med	ND	Low	
C ₇ H ₁₄ O ₆ S	ND	Low	ND	Low	
C ₈ H ₁₂ O ₇ S	ND	Low	ND	ND	
C ₈ H ₁₂ O ₈ S	Med	Low	ND	Low	
C ₈ H ₁₄ O ₆ S	Low	Med	ND	Low	
C ₈ H ₁₄ O ₇ S	ND	Low	ND	ND	
C ₈ H ₁₆ O ₆ S	Med	Med	Low	Med	
C ₉ H ₁₂ O ₃	Low	Low	ND	Low	methyl-syringol
C ₉ H ₁₂ O ₅	Med	Med	Low	Low	
C ₉ H ₁₂ O ₇ S	Low	ND	ND	ND	
C ₉ H ₁₄ O ₆ S	Med	ND	ND	ND	
C ₉ H ₁₄ O ₇ S	Low	Med	ND	Low	
C ₉ H ₁₄ O ₈ S	Med	Med	ND	Low	
C ₉ H ₁₄ O ₉ S	Low	Low	ND	ND	
C ₉ H ₁₅ NO ₈ S	Very High	Very High	Low	Very High	
C ₉ H ₁₆ O ₆ S	Med	Med	ND	Med	
C ₉ H ₁₆ O ₇ S	Low	Med	ND	Low	
C ₉ H ₁₆ O ₈ S	ND	Low	ND	ND	
C ₉ H ₁₈ O ₆ S	Med	Med	Low	Med	
C ₉ H ₁₈ O ₈ S	ND	ND	ND	Low	
C ₁₀ H ₁₄ O ₅	Med	Med	Low	Low	
C ₁₀ H ₁₄ O ₆	Med	Med	ND	Low	
C ₁₀ H ₁₄ O ₇ S	Med	ND	ND	ND	
C ₁₀ H ₁₄ O ₈ S	Low	Low	ND	ND	
C ₁₀ H ₁₆ O ₆ S	Med	ND	ND	ND	
C ₁₀ H ₁₆ O ₇ S	Med	High	ND	Low	
C ₁₀ H ₁₆ O ₈ S	ND	Med	ND	Low	
C ₁₀ H ₁₆ O ₉ S	Low	Low	ND	ND	
C ₁₀ H ₁₇ NO ₇ S	Very High	Very High	High	Very High	
C ₁₀ H ₁₇ NO ₁₀ S	Med	Med	Low	Med	
C ₁₀ H ₁₈ O ₅ S	ND	Med	ND	ND	
C ₁₀ H ₁₈ O ₇ S	Med	Med	ND	Low	

Formula^a	SPC0106F^b	SPC0201F^b	BO0204N^b	BO0213D^b	Possible Identity
C ₁₀ H ₁₈ O ₈ S	Low	Med	ND	ND	
C ₁₀ H ₁₉ NO ₉ S	High	Med	Low	Med	
C ₁₀ H ₂₀ O ₅ S	Med	Med	Low	High	
C ₁₁ H ₂₁ NO ₉ S	Med	Med	Low	Med	
C ₁₁ H ₂₂ O ₅ S	Med	Med	Low	High	
C ₁₂ H ₂₀ O ₇ S	Med	Med	ND	Med	
C ₁₂ H ₂₂ O ₇ S	Med	Med	Low	Med	
C ₁₂ H ₂₃ NO ₉ S	Med	Med	Low	Med	
C ₁₂ H ₂₄ O ₅ S	Med	Low	Low	High	
C ₁₂ H ₂₆ O ₄ S	High	ND	Med	Low	
C ₁₃ H ₂₄ O ₇ S	Med	Med	Low	Med	
C ₁₃ H ₂₆ O ₆ S	Med	Low	Low	Med	
C ₁₄ H ₂₀ O ₉	Low	Low	ND	Med	
C ₁₄ H ₂₄ O ₈	ND	Low	ND	ND	
C ₁₄ H ₂₇ NO ₉ S	Med	Low	Low	Med	
C ₁₄ H ₂₈ O ₅ S	Med	Low	Low	Med	
C ₁₄ H ₂₈ O ₆ S	Med	Low	Low	Med	
C ₁₄ H ₃₀ O ₄ S	Med	ND	Low	Med	
C ₁₅ H ₂₄ O ₉	Low	Low	ND	Low	
C ₁₅ H ₂₉ NO ₉ S	Med	Low	Low	Med	
C ₁₅ H ₃₀ O ₆ S	Med	Low	Low	Med	
C ₁₆ H ₂₄ O ₁₁ S	ND	Low	ND	ND	
C ₁₆ H ₃₁ NO ₉ S	Med	Low	Low	Med	
C ₁₇ H ₃₃ NO ₉ S	Med	Low	Low	Med	
C ₁₈ H ₂₆ O ₁₂ S	ND	Low	ND	ND	
C ₁₈ H ₂₈ O ₁₁ S	Low	Low	ND	ND	
C ₁₈ H ₃₈ O ₆ S	Low	ND	ND	Low	
C ₁₉ H ₃₀ O ₁₂ S	Low	ND	ND	ND	

^aAs the identified molecular formulas likely represent a variety of structural isomers, we note that matched molecular formulas do not necessarily correspond to the same molecular structure or atmospheric origin. ^bThe normalized abundances are indicated for each sample, with: “ND” (not detected), “Low” ($\leq 3\%$), “Med”, ($> 3\%$ and $\leq 15\%$), “High” ($> 15\%$ and $\leq 50\%$) and “Very High” ($> 50\%$).

6.2 Ancillary figures

Figure 6.1. Reconstructed mass spectra for SPC01 daytime aerosol sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.2. Reconstructed mass spectra for SPC01 nighttime aerosol sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.3. Reconstructed mass spectra for SPC01 fog water sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.4. Reconstructed mass spectra for SPC02 daytime aerosol sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.5. Reconstructed mass spectra for SPC02 nighttime aerosol sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.6. Reconstructed mass spectra for SPC02 fog water sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.7. Reconstructed mass spectra for SPC03 daytime aerosol sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.8. Reconstructed mass spectra for SPC03 nighttime aerosol sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.9. Reconstructed mass spectra for SPC03 fog water sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.10. Reconstructed mass spectra for SPC04 daytime aerosol sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.11. Reconstructed mass spectra for SPC04 fog water sample collected by -ESI (upper) and +ESI (lower). Only masses assigned a molecular formula are shown. This mass spectrum is the result of the merger between two scan ranges (m/z 100 to 800 and m/z 200 to 800) collected for this sample, as described in section 3.3.

Figure 6.12. Total ion chromatogram for the liquid chromatography-mass spectrometry analysis of Suwannee River Fulvic Acid (SRFA), a reference complex organic mixture from the International Humic Substances Society. Replicate analyses (“SRFA_r1”, “SRFA_r2” and “SRFA_r3” are shown, with excellent agreement. Samples were analyzed with a step-gradient profile: beginning with 5% acetonitrile in water (0.1% formic acid) for 5 minutes, then increasing to 20% acetonitrile for 10 minutes, 40% acetonitrile for 10 minutes, 60% acetonitrile for 10 minutes and 90% acetonitrile for 10 minutes. Samples of 5 μL of 500 $\text{ng } \mu\text{L}^{-1}$ SRFA were injected onto an Agilent PLRP-S (100 $\text{\AA}$, 3 μm 150 x 2.1 mm) analytical column and analyzed by -ESI using an Orbitrap Elite mass spectrometer at a resolving power of 240,000. The injected samples separated into two broad peaks (~ 10 min and ~ 20 min) over the 45 min analysis. The total ion chromatogram was separated into 22, 2 min fractions (f1, f2, f3... etc.) for molecular formula assignment using Composer (Sierra Analytics v. 1.0.5). The comparison of molecular formulas from a select sub-set of fractions (f3, f5, f10 and f15; indicated with blue bars) is shown in Fig. 6.13. Overall, $\sim 20,000$ molecular formulas were assigned to the total ion chromatogram for SRFA with this method.

Figure 6.13. Kendrick mass defect vs. Kendrick mass for the molecular formula assignment of a sub-set of sample fractions (fx) for Suwannee River Fulvic Acid (SRFA) analyzed by liquid chromatography-mass spectrometry (see Fig. 6.12). Formulas of the O₁₀ sub-class are shown (all molecular formulas have 10 oxygen atoms and some other combination of carbon and hydrogen). The molecular formulas of a fraction are shown in each row, with increasing fraction number from top to bottom. From left to right the formulas of these fractions are compared to the other fractions (i.e. “f3 & f5” shows the formulas in f3 that are also present in f5). Intersecting formulas are plotted with color scaled log₁₀ intensity values of the compared fraction (i.e. in “f3 & f5” the formulas are color scaled to the log₁₀ intensity of f3, where in “f5 & f3” the formulas are color scaled to the log₁₀ intensity of f5). The remaining formulas are plotted in gray. Approximately 20,000 molecular formulas were assigned to the 22 fractions of this analysis of SRFA, where many these formulas are expected to be structural isomers (identical formulas appearing in different fractions).

7 Appendix B: copyright documentation

The content of chapter 2 was previously published as an article in the journal *Atmospheric Chemistry and Physics*:

Brege, M., Paglione, M., Gilardoni, S., Decesari, S., Facchini, M.C. and Mazzoleni, L.R. (2018). "Molecular insights on aging and aqueous-phase processing from ambient biomass burning emissions-influenced po valley fog and aerosol." *Atmospheric Chemistry and Physics* 18(17): 13197-13214.

As this article is distributed under the Creative Commons Attribution 4.0 License, the authors retain the copyright of the article, as stated in the copyright page of the *Atmospheric Chemistry and Physics* website: https://www.atmospheric-chemistry-and-physics.net/about/licence_and_copyright.html (accessed 15-Oct 2019). Under the Creative Commons Attribution 4.0 License, the published material may be shared and adapted with appropriate citation and credit given to the original authors. As the publisher does not require further permission for the reproduction of articles, the content may be republished with permission from the authors of the article. Sufficient citation of the original work has been provided and as the lead author of the article I grant permission for its content to be republished here. The Creative Commons Attribution 4.0 License can be viewed at: <https://creativecommons.org/licenses/by/4.0/> (accessed 15-Oct 2019).

8 References

- 1) Ackerman, A. S., Toon, O. B., Stevens, D. E., Heymsfield, A. J., Ramanathan, V. and Welton, E. J. (2000). "Reduction of tropical cloudiness by soot." *Science* 288(5468): 1042-1047.
- 2) Adachi, K., Sedlacek, A. J., Kleinman, L., Springston, S. R., Wang, J., Chand, D., Hubbe, J. M., Shilling, J. E., Onasch, T. B. and Kinase, T. J. P. o. t. N. A. o. S. (2019). "Spherical tarball particles form through rapid chemical and physical changes of organic matter in biomass-burning smoke." 201900129.
- 3) Aiken, A. C., Decarlo, P. F., Kroll, J. H., Worsnop, D. R., Huffman, J. A., Docherty, K. S., Ulbrich, I. M., Mohr, C., Kimmel, J. R., Sueper, D., Sun, Y., Zhang, Q., Trimborn, A., Northway, M., Ziemann, P. J., Canagaratna, M. R., Onasch, T. B., Alfarra, M. R., Prevot, A. S. H., Dommen, J., Duplissy, J., Metzger, A., Baltensperger, U. and Jimenez, J. L. (2008). "O/c and om/oc ratios of primary, secondary, and ambient organic aerosols with high-resolution time-of-flight aerosol mass spectrometry." *Environmental Science & Technology* 42(12): 4478-4485.
- 4) Alexander, D. T. L., Crozier, P. A. and Anderson, J. R. (2008). "Brown carbon spheres in east asian outflow and their optical properties." *Science* 321(5890): 833-836.
- 5) Altieri, K. E., Turpin, B. J. and Seitzinger, S. P. (2009). "Oligomers, organosulfates, and nitrooxy organosulfates in rainwater identified by ultra-high resolution electrospray ionization ft-icr mass spectrometry." *Atmospheric Chemistry and Physics* 9(7): 2533-2542.
- 6) Alves, C. A., Vicente, A. M., Custodio, D., Cerqueira, M., Nunes, T., Pio, C., Lucarelli, F., Calzolari, G., Nava, S., Diapouli, E., Eleftheriadis, K., Querol, X. and Bandowe, B. A. M. (2017). "Polycyclic aromatic hydrocarbons and their derivatives (nitro-pahs, oxygenated pahs, and azaarenes) in pm2.5 from southern european cities." *Science of the Total Environment* 595: 494-504.
- 7) Anderson, T. R., Hawkins, E. and Jones, P. D. (2016). "Co2, the greenhouse effect and global warming: From the pioneering work of arrhenius and callendar to today's earth system models." *Endeavour* 40(3): 178-187.
- 8) Andreae, M. O. (2001). "The dark side of aerosols." *Nature* 409(6821): 671-672.
- 9) Andreae, M. O. and Gelencser, A. (2006). "Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols." *Atmospheric Chemistry and Physics* 6: 3131-3148.
- 10) Andreae, M. O. and Merlet, P. (2001). "Emission of trace gases and aerosols from biomass burning." *Global Biogeochemical Cycles* 15(4): 955-966.

- 11) Bahreini, R., Ervens, B., Middlebrook, A. M., Warneke, C., de Gouw, J. A., DeCarlo, P. F., Jimenez, J. L., Brock, C. A., Neuman, J. A., Ryerson, T. B., Stark, H., Atlas, E., Brioude, J., Fried, A., Holloway, J. S., Peischl, J., Richter, D., Walega, J., Weibring, P., Wollny, A. G. and Fehsenfeld, F. C. (2009). "Organic aerosol formation in urban and industrial plumes near houston and dallas, texas." *Journal of Geophysical Research-Atmospheres* 114.
- 12) Barbero, R., Abatzoglou, J. T., Larkin, N. K., Kolden, C. A. and Stocks, B. (2015). "Climate change presents increased potential for very large fires in the contiguous united states." *International Journal of Wildland Fire* 24(7): 892-899.
- 13) Bennett, J. E., Tamura-Wicks, H., Parks, R. M., Burnett, R. T., Pope, C. A., Bechle, M. J., Marshall, J. D., Danaei, G. and Ezzati, M. (2019). "Particulate matter air pollution and national and county life expectancy loss in the USA: A spatiotemporal analysis." *Plos Medicine* 16(7).
- 14) Bhandari, J. (2018). "Morphology and mixing state of soot and tar balls: Implications for optical properties and climate." *Open Access Dissertation, Michigan Technological University*.
- 15) Bianchi, F., Trostl, J., Junninen, H., Frege, C., Henne, S., Hoyle, C. R., Molteni, U., Herrmann, E., Adamov, A., Bukowiecki, N., Chen, X., Duplissy, J., Gysel, M., Hutterli, M., Kangasluoma, J., Kontkanen, J., Kurten, A., Manninen, H. E., Munch, S., Perakyla, O., Petaja, T., Rondo, L., Williamson, C., Weingartner, E., Curtius, J., Worsnop, D. R., Kulmala, M., Dommen, J. and Baltensperger, U. (2016). "New particle formation in the free troposphere: A question of chemistry and timing." *Science* 352(6289): 1109-1112.
- 16) Bond, T. C., Streets, D. G., Yarber, K. F., Nelson, S. M., Woo, J. H. and Klimont, Z. (2004). "A technology-based global inventory of black and organic carbon emissions from combustion." *Journal of Geophysical Research-Atmospheres* 109(D14): D14203.
- 17) Boone, E. J., Laskin, A., Laskin, J., Wirth, C., Shepson, P. B., Stirm, B. H. and Pratt, K. A. (2015). "Aqueous processing of atmospheric organic particles in cloud water collected via aircraft sampling." *Environmental Science & Technology* 49(14): 8523-8530.
- 18) Boyd, C. M., Sanchez, J., Xu, L., Eugene, A. J., Nah, T., Tuet, W. Y., Guzman, M. I. and Ng, N. L. (2015). "Secondary organic aerosol formation from the beta-pinene+no₃ system: Effect of humidity and peroxy radical fate." *Atmospheric Chemistry and Physics* 15(13): 7497-7522.
- 19) Brege, M., Paglione, M., Gilardoni, S., Decesari, S., Facchini, M. C. and Mazzoleni, L. R. (2018). "Molecular insights on aging and aqueous-phase processing from

- ambient biomass burning emissions-influenced po valley fog and aerosol." *Atmospheric Chemistry and Physics* 18(17): 13197-13214.
- 20) Cech, N. B. and Enke, C. G. (2001). "Practical implications of some recent studies in electrospray ionization fundamentals." *Mass Spectrometry Reviews* 20(6): 362-387.
 - 21) Chakrabarty, R. K., Moosmuller, H., Chen, L. W. A., Lewis, K., Arnott, W. P., Mazzoleni, C., Dubey, M. K., Wold, C. E., Hao, W. M. and Kreidenweis, S. M. (2010). "Brown carbon in tar balls from smoldering biomass combustion." *Atmospheric Chemistry and Physics* 10(13): 6363-6370.
 - 22) Chandrakar, K. K., Cantrell, W., Chang, K., Ciochetto, D., Niedermeier, D., Ovchinnikov, M., Shaw, R. A. and Yang, F. (2016). "Aerosol indirect effect from turbulence-induced broadening of cloud-droplet size distributions." *Proceedings of the National Academy of Sciences of the United States of America* 113(50): 14243-14248.
 - 23) Chang, J. L. and Thompson, J. E. (2010). "Characterization of colored products formed during irradiation of aqueous solutions containing h₂o₂ and phenolic compounds." *Atmospheric Environment* 44(4): 541-551.
 - 24) China, S., Mazzoleni, C., Gorkowski, K., Aiken, A. C. and Dubey, M. K. (2013). "Morphology and mixing state of individual freshly emitted wildfire carbonaceous particles." *Nature Communications* 4.
 - 25) Chow, J. C., Watson, J. G., Edgerton, S. A. and Vega, E. (2002). "Chemical composition of pm_{2.5} and pm₁₀ in mexico city during winter 1997." *Science of the Total Environment* 287(3): 177-201.
 - 26) Ciobanu, V. G., Marcolli, C., Krieger, U. K., Zuend, A. and Peter, T. (2010). "Efflorescence of ammonium sulfate and coated ammonium sulfate particles: Evidence for surface nucleation." *Journal of Physical Chemistry A* 114(35): 9486-9495.
 - 27) Cook, R. D., Lin, Y. H., Peng, Z. Y., Boone, E., Chu, R. K., Dukett, J. E., Gunsch, M. J., Zhang, W. L., Tolic, N., Laskin, A. and Pratt, K. A. (2017). "Biogenic, urban, and wildfire influences on the molecular composition of dissolved organic compounds in cloud water." *Atmospheric Chemistry and Physics* 17(24): 15167-15180.
 - 28) Costabile, F., Alas, H., Aufderheide, M., Avino, P., Amato, F., Argentini, S., Barnaba, F., Berico, M., Bernardoni, V., Biondi, R., Calzolari, G., Canepari, S., Casasanta, G., Ciampichetti, S., Conidi, A., Cordelli, E., Di Ianni, A., Di Liberto, L., Facchini, M. C., Facci, A., Frasca, D., Gilardoni, S., Grollino, M. G., Gualtieri, M., Lucarelli, F., Malaguti, A., Manigrasso, M., Montagnoli, M., Nava,

- S., Padoan, E., Perrino, C., Petralia, E., Petenko, I., Querol, X., Simonetti, G., Tranfo, G., Ubertini, S., Valli, G., Valentini, S., Vecchi, R., Volpi, F., Weinhold, K., Wiedensohler, A., Zanini, G. and Gobbi, G. P. (2017). "First results of the "carbonaceous aerosol in rome and environs (care)" experiment: Beyond current standards for pm10." *Atmosphere* 8(12).
- 29) Costabile, F., Gilardoni, S., Barnaba, F., Di Ianni, A., Di Liberto, L., Dionisi, D., Manigrasso, M., Paglione, M., Poluzzi, V., Rinaldi, M., Facchini, M. C. and Gobbi, G. P. (2017). "Characteristics of brown carbon in the urban po valley atmosphere." *Atmospheric Chemistry and Physics* 17(1): 313-326.
- 30) Cubison, M. J., Ortega, A. M., Hayes, P. L., Farmer, D. K., Day, D., Lechner, M. J., Brune, W. H., Apel, E., Diskin, G. S., Fisher, J. A., Fuelberg, H. E., Hecobian, A., Knapp, D. J., Mikoviny, T., Riemer, D., Sachse, G. W., Sessions, W., Weber, R. J., Weinheimer, A. J., Wisthaler, A. and Jimenez, J. L. (2011). "Effects of aging on organic aerosol from open biomass burning smoke in aircraft and laboratory studies." *Atmospheric Chemistry and Physics* 11(23): 12049-12064.
- 31) Dall'Osto, M., Paglione, M., Decesari, S., Facchini, M. C., O'Dowd, C., Plass-Duellmer, C. and Harrison, R. M. (2015). "On the origin of ams "cooking organic aerosol" at a rural site." *Environmental Science & Technology* 49(24): 13964-13972.
- 32) Damoah, R., Spichtinger, N., Forster, C., James, P., Mattis, I., Wandinger, U., Beirle, S., Wagner, T. and Stohl, A. (2004). "Around the world in 17 days - hemispheric-scale transport of forest fire smoke from russia in may 2003." *Atmospheric Chemistry and Physics* 4: 1311-1321.
- 33) Darer, A. I., Cole-Filipiak, N. C., O'Connor, A. E. and Elrod, M. J. (2011). "Formation and stability of atmospherically relevant isoprene-derived organosulfates and organonitrates." *Environmental Science & Technology* 45(5): 1895-1902.
- 34) De Haan, D. O., Hawkins, L. N., Kononenko, J. A., Turley, J. J., Corrigan, A. L., Tolbert, M. A. and 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.
- 35) DeCarlo, P. F., Kimmel, J. R., Trimborn, A., Northway, M. J., Jayne, J. T., Aiken, A. C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K. S., Worsnop, D. R. and Jimenez, J. L. (2006). "Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer." *Analytical Chemistry* 78(24): 8281-8289.
- 36) Decesari, S., Facchini, M. C., Fuzzi, S. and Tagliavini, E. (2000). "Characterization of water-soluble organic compounds in atmospheric aerosol: A new approach." *Journal of Geophysical Research-Atmospheres* 105(D1): 1481-1489.

- 37) Decesari, S., Mircea, M., Cavalli, F., Fuzzi, S., Moretti, F., Tagliavini, E. and Facchini, M. C. (2007). "Source attribution of water-soluble organic aerosol by nuclear magnetic resonance spectroscopy." *Environmental Science & Technology* 41(7): 2479-2484.
- 38) Deguillaume, L., Leriche, M. and Chaurnerliac, N. (2005). "Impact of radical versus non-radical pathway in the fenton chemistry on the iron redox cycle in clouds." *Chemosphere* 60(5): 718-724.
- 39) Denisov, E., Damoc, E., Lange, O. and Makarov, A. (2012). "Orbitrap mass spectrometry with resolving powers above 1,000,000." *International Journal of Mass Spectrometry* 325: 80-85.
- 40) DeRieux, W. S., Li, Y., Lin, P., Laskin, J., Laskin, A., Bertram, A. K., Nizkorodov, S. A. and Shiraiwa, M. (2018). "Predicting the glass transition temperature and viscosity of secondary organic material using molecular composition." *Atmospheric Chemistry and Physics* 18(9): 6331-6351.
- 41) Despres, V. R., Huffman, J. A., Burrows, S. M., Hoose, C., Safatov, A. S., Buryak, G., Frohlich-Nowoisky, J., Elbert, W., Andreae, M. O., Poschl, U. and Jaenicke, R. (2012). "Primary biological aerosol particles in the atmosphere: A review." *Tellus Series B-Chemical and Physical Meteorology* 64.
- 42) Desyaterik, Y., Sun, Y., Shen, X. H., Lee, T. Y., Wang, X. F., Wang, T. and Collett, J. L. (2013). "Speciation of "brown" carbon in cloud water impacted by agricultural biomass burning in eastern china." *Journal of Geophysical Research-Atmospheres* 118(13): 7389-7399.
- 43) Dickhut, R. M., Andren, A. W. and Armstrong, D. E. (1989). "Naphthalene solubility in selected organic solvent-water mixtures." *Journal of Chemical and Engineering Data* 34(4): 438-443.
- 44) Drewnick, F., Hings, S. S., DeCarlo, P., Jayne, J. T., Gonin, M., Fuhrer, K., Weimer, S., Jimenez, J. L., Demerjian, K. L., Borrmann, S. and Worsnop, D. R. (2005). "A new time-of-flight aerosol mass spectrometer (tof-ams) - instrument description and first field deployment." *Aerosol Science and Technology* 39(7): 637-658.
- 45) Dzepina, K., Mazzoleni, C., Fialho, P., China, S., Zhang, B., Owen, R. C., Helmig, D., Hueber, J., Kumar, S., Perlinger, J. A., Kramer, L. J., Dziobak, M. P., Ampadu, M. T., Olsen, S., Wuebbles, D. J. and Mazzoleni, L. R. (2015). "Molecular characterization of free tropospheric aerosol collected at the pico mountain observatory: A case study with a long-range transported biomass burning plume." *Atmospheric Chemistry and Physics* 15(9): 5047-5068.

- 46) El-Sayed, M. M. H., Wang, Y. Q. and Hennigan, C. J. (2015). "Direct atmospheric evidence for the irreversible formation of aqueous secondary organic aerosol." *Geophysical Research Letters* 42(13): 5577-5586.
- 47) Ervens, B. (2015). "Modeling the processing of aerosol and trace gases in clouds and fogs." *Chemical Reviews* 115(10): 4157-4198.
- 48) Ervens, B., Turpin, B. J. and Weber, R. J. (2011). "Secondary organic aerosol formation in cloud droplets and aqueous particles (aqsoa): A review of laboratory, field and model studies." *Atmospheric Chemistry and Physics* 11(21): 11069-11102.
- 49) Ervens, B., Wang, Y., Eagar, J., Leaitch, W. R., Macdonald, A. M., Valsaraj, K. T. and Herckes, P. (2013). "Dissolved organic carbon (doc) and select aldehydes in cloud and fog water: The role of the aqueous phase in impacting trace gas budgets." *Atmospheric Chemistry and Physics* 13(10): 5117-5135.
- 50) Facchini, M. C., Fuzzi, S., Zappoli, S., Andracchio, A., Gelencser, A., Kiss, G., Krivacsy, Z., Meszaros, E., Hansson, H. C., Alsberg, T. and Zebuhr, Y. (1999). "Partitioning of the organic aerosol component between fog droplets and interstitial air." *Journal of Geophysical Research-Atmospheres* 104(D21): 26821-26832.
- 51) Fan, C. H. and Jafvert, C. T. (1997). "Margules equations applied to pah solubilities in alcohol-water mixtures." *Environmental Science & Technology* 31(12): 3516-3522.
- 52) Farmer, P. B. (2003). "Molecular epidemiology studies of carcinogenic environmental pollutants - effects of polycyclic aromatic hydrocarbons (pahs) in environmental pollution on exogenous and oxidative DNA damage." *Mutation Research-Reviews in Mutation Research* 544(2-3): 397-402.
- 53) Feingold, G., Eberhard, W. L., Veron, D. E. and Previdi, M. (2003). "First measurements of the twomey indirect effect using ground-based remote sensors." *Geophysical Research Letters* 30(6).
- 54) Fialho, P., Hansen, A. D. A. and Honrath, R. E. (2005). "Absorption coefficients by aerosols in remote areas: A new approach to decouple dust and black carbon absorption coefficients using seven-wavelength aethalometer data." *Journal of Aerosol Science* 36(2): 267-282.
- 55) Fleming, L. T., Lin, P., Laskin, A., Laskin, J., Weltman, R., Edwards, R. D., Arora, N. K., Yadav, A., Meinardi, S., Blake, D. R., Pillarisetti, A., Smith, K. R. and Nizkorodov, S. A. (2018). "Molecular composition of particulate matter emissions from dung and brushwood burning household cookstoves in haryana, india." *Atmospheric Chemistry and Physics* 18(4): 2461-2480.

- 56) Fleming, L. T., Weltman, R., Yadav, A., Edwards, R. D., Arora, N. K., Pillarisetti, A., Meinardi, S., Smith, K. R., Blake, D. R., Nizkorodov, S. A. J. A. C. and Physics (2018). "Emissions from village cookstoves in haryana, india, and their potential impacts on air quality." 18(20): 15169-15182.
- 57) Fuzzi, S., Baltensperger, U., Carslaw, K., Decesari, S., van Der Gon, H. D., Facchini, M. C., Fowler, D., Koren, I., Langford, B., Lohmann, U., Nemitz, E., Pandis, S., Riipinen, I., Rudich, Y., Schaap, M., Slowik, J. G., Spracklen, D. V., Vignati, E., Wild, M., Williams, M. and Gilardoni, S. (2015). "Particulate matter, air quality and climate: Lessons learned and future needs." *Atmospheric Chemistry and Physics* 15(14): 8217-8299.
- 58) Fuzzi, S., Orsi, G., Bonforte, G., Zardini, B. and Franchini, P. L. (1997). "An automated fog water collector suitable for deposition networks: Design, operation and field tests." *Water Air and Soil Pollution* 93(1-4): 383-394.
- 59) Galloway, M. M., Powelson, M. H., Sedehi, N., Wood, S. E., Millage, K. D., Kononenko, J. A., Rynaski, A. D. and De Haan, D. O. (2014). "Secondary organic aerosol formation during evaporation of droplets containing atmospheric aldehydes, amines, and ammonium sulfate." *Environmental science & technology* 48(24): 14417-14425.
- 60) Gavard, R., Palacio Lozano, D. C., Guzman, A., Rossell, D., Spencer, S. E. and Barrow, M. P. J. A. c. (2019). "Rhapso: Automatic stitching of mass segments from fourier transform ion cyclotron resonance mass spectra."
- 61) Gelencsér, A. (2005). Carbonaceous aerosol, Springer Science & Business Media.
- 62) Gergel, D. R., Nijssen, B., Abatzoglou, J. T., Lettenmaier, D. P. and Stumbaugh, M. R. (2017). "Effects of climate change on snowpack and fire potential in the western USA." *Climatic Change* 141(2): 287-299.
- 63) Gilardoni, S., Massoli, P., Giulianelli, L., Rinaldi, M., Paglione, M., Pollini, F., Lanconelli, C., Poluzzi, V., Carbone, S., Hillamo, R., Russell, L. M., Facchini, M. C. and Fuzzi, S. (2014). "Fog scavenging of organic and inorganic aerosol in the po valley." *Atmospheric Chemistry and Physics* 14(13): 6967-6981.
- 64) Gilardoni, S., Massoli, P., Paglione, M., Giulianelli, L., Carbone, C., Rinaldi, M., Decesari, S., Sandrini, S., Costabile, F., Gobbi, G. P., Pietrogrande, M. C., Visentin, M., Scotto, F., Fuzzi, S. and Facchini, M. C. (2016). "Direct observation of aqueous secondary organic aerosol from biomass-burning emissions." *Proceedings of the National Academy of Sciences of the United States of America* 113(36): 10013-10018.
- 65) Giulianelli, L., Gilardoni, S., Tarozzi, L., Rinaldi, M., Decesari, S., Carbone, C., Facchini, M. C. and Fuzzi, S. (2014). "Fog occurrence and chemical composition

- in the po valley over the last twenty years." *Atmospheric Environment* 98: 394-401.
- 66) Glasius, M., Ketzel, M., Wahlin, P., Jensen, B., Monster, J., Berkowicz, R. and Palmgren, F. (2006). "Impact of wood combustion on particle levels in a residential area in denmark." *Atmospheric Environment* 40(37): 7115-7124.
 - 67) Goldstein, A. H. and Galbally, I. E. (2007). "Known and unexplored organic constituents in the earth's atmosphere." *Environmental Science & Technology* 41(5): 1514-1521.
 - 68) Gross, J. H. (2004). Mass spectrometry: A textbook, Springer-Verlag.
 - 69) Han, T. T., Liu, X. G., Zhang, Y. H., Qu, Y., Zeng, L. M., Hu, M. and Zhu, T. (2015). "Role of secondary aerosols in haze formation in summer in the megacity beijing." *Journal of Environmental Sciences-China* 31: 51-60.
 - 70) Hand, J. L., Malm, W. C., Laskin, A., Day, D., Lee, T., Wang, C., Carrico, C., Carrillo, J., Cowin, J. P., Collett, J. and Iedema, M. J. (2005). "Optical, physical, and chemical properties of tar balls observed during the yosemite aerosol characterization study." *Journal of Geophysical Research-Atmospheres* 110(D21): D21210.
 - 71) Harrison, M. A., Barra, S., Borghesi, D., Vione, D., Arsene, C. and Olariu, R. I. J. A. E. (2005). "Nitrated phenols in the atmosphere: A review." 39(2): 231-248.
 - 72) Harrison, M. A. J., Barra, S., Borghesi, D., Vione, D., Arsene, C. and Olariu, R. L. (2005). "Nitrated phenols in the atmosphere: A review." *Atmospheric Environment* 39(2): 231-248.
 - 73) Hawkes, J. A., Patriarca, C., Sjoberg, P. J. R., Tranvik, L. and Bergquist, J. (2018). "Extreme isomeric complexity of dissolved organic matter found across aquatic environments." *Limnology and Oceanography Letters* 3(2): 21-30.
 - 74) Hawkins, L., Welsh, H. and Alexander, M. (2018). "Evidence for pyrazine-based chromophores in cloudwater mimics containing methylglyoxal and ammonium sulfate." *Atmos. Chem. Phys. Discuss.*: 29.
 - 75) Hawkins, L. N., Lemire, A. N., Galloway, M. M., Corrigan, A. L., Turley, J. J., Espelien, B. M. and De Haan, D. O. (2016). "Maillard chemistry in clouds and aqueous aerosol as a source of atmospheric humic-like substances." *Environmental Science & Technology* 50(14): 7443-7452.
 - 76) He, Q. F., Ding, X., Wang, X. M., Yu, J. Z., Fu, X. X., Liu, T. Y., Zhang, Z., Xue, J., Chen, D. H., Zhong, L. J. and Donahue, N. M. (2014). "Organosulfates from pinene and isoprene over the pearl river delta, south china: Seasonal variation and

- implication in formation mechanisms." *Environmental Science & Technology* 48(16): 9236-9245.
- 77) Herckes, P., Chang, H., Lee, T. and Collett, J. L. (2007). "Air pollution processing by radiation fogs." *Water Air and Soil Pollution* 181(1-4): 65-75.
- 78) Herckes, P., Marcotte, A. R., Wang, Y. and Collett, J. L. (2015). "Fog composition in the central valley of california over three decades." *Atmospheric Research* 151: 20-30.
- 79) Herckes, P., Valsaraj, K. T. and Collett, J. L. (2013). "A review of observations of organic matter in fogs and clouds: Origin, processing and fate." *Atmospheric Research* 132: 434-449.
- 80) Herrmann, H., Schaefer, T., Tilgner, A., Styler, S. A., Weller, C., Teich, M. and Otto, T. (2015). "Tropospheric aqueous-phase chemistry: Kinetics, mechanisms, and its coupling to a changing gas phase." *Chemical Reviews* 115(10): 4259-4334.
- 81) Hertkorn, N., Ruecker, C., Meringer, M., Gugisch, R., Frommberger, M., Perdue, E. M., Witt, M. and Schmitt-Kopplin, P. (2007). "High-precision frequency measurements: Indispensable tools at the core of the molecular-level analysis of complex systems." *Analytical and Bioanalytical Chemistry* 389(5): 1311-1327.
- 82) Herzsprung, P., Hertkorn, N., von Tumpling, W., Harir, M., Frieze, K. and 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." *Analytical and Bioanalytical Chemistry* 406(30): 7977-7987.
- 83) Hinds, W. C. (1999). Aerosol technology: Properties, behavior, and measurement of airborne particles, John Wiley & Sons.
- 84) Hinks, M. L., Brady, M. V., Lignell, H., Song, M. J., Grayson, J. W., Bertram, A. K., Lin, P., Laskin, A., Laskin, J. and Nizkorodov, S. A. (2016). "Effect of viscosity on photodegradation rates in complex secondary organic aerosol materials." *Physical Chemistry Chemical Physics* 18(13): 8785-8793.
- 85) Hoffman, E. d. and Stroobant, V. (2007). Mass spectrometry: Principles and applications, John Wiley & Sons Inc.
- 86) Holmes, B. J. and Petrucci, G. A. (2007). "Oligomerization of levoglucosan by fenton chemistry in proxies of biomass burning aerosols." *Journal of Atmospheric Chemistry* 58(2): 151-166.

- 87) Hu, K. S., Darer, A. I. and Elrod, M. J. (2011). "Thermodynamics and kinetics of the hydrolysis of atmospherically relevant organonitrates and organosulfates." *Atmospheric Chemistry and Physics* 11(16): 8307-8320.
- 88) Huang, D. D., Zhang, Q., Cheung, H. H. Y., Yu, L., Zhou, S., Anastasio, C., Smith, J. D. and Chan, C. K. (2018). "Formation and evolution of aqsoa from aqueous-phase reactions of phenolic carbonyls: Comparison between ammonium sulfate and ammonium nitrate solutions." *Environmental Science & Technology* 52(16): 9215-9224.
- 89) Hughey, C. A., Hendrickson, C. L., Rodgers, R. P., Marshall, A. G. and Qian, K. N. (2001). "Kendrick mass defect spectrum: A compact visual analysis for ultrahigh-resolution broadband mass spectra." *Analytical Chemistry* 73(19): 4676-4681.
- 90) IPCC (2018). Global warming of 1.5° c: An ipcc special report on the impacts of global warming of 1.5° c above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, Intergovernmental Panel on Climate Change.
- 91) Jacobson, M. Z. (2002). "Control of fossil-fuel particulate black carbon and organic matter, possibly the most effective method of slowing global warming." *Journal of Geophysical Research-Atmospheres* 107(D19).
- 92) Kampf, C. J., Waxman, E. M., Slowik, J. G., Dommen, J., Pfaffenberger, L., Praplan, A. P., Prevot, A. S. H., Baltensperger, U., Hoffmann, T. and Volkamer, R. (2013). "Effective henry's law partitioning and the salting constant of glyoxal in aerosols containing sulfate." *Environmental Science & Technology* 47(9): 4236-4244.
- 93) Kanakidou, M., Seinfeld, J. H., Pandis, S. N., Barnes, I., Dentener, F. J., Facchini, M. C., Van Dingenen, R., Ervens, B., Nenes, A., Nielsen, C. J., Swietlicki, E., Putaud, J. P., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G. K., Winterhalter, R., Myhre, C. E. L., Tsigaridis, K., Vignati, E., Stephanou, E. G. and Wilson, J. (2005). "Organic aerosol and global climate modelling: A review." *Atmospheric Chemistry and Physics* 5: 1053-1123.
- 94) Kauppila, T. J., Kuuranne, T., Meurer, E. C., Eberlin, M. N., Kotiaho, T. and Kostiainen, R. (2002). "Atmospheric pressure photoionization mass spectrometry. Ionization mechanism and the effect of solvent on the ionization of naphthalenes." *Analytical Chemistry* 74(21): 5470-5479.
- 95) Khain, A., Ovtchinnikov, M., Pinsky, M., Pokrovsky, A. and Krugliak, H. (2000). "Notes on the state-of-the-art numerical modeling of cloud microphysics." *Atmospheric Research* 55(3-4): 159-224.

- 96) Kitanovski, Z., Grgic, I., Vermeylen, R., Claeys, M. and Maenhaut, W. (2012). "Liquid chromatography tandem mass spectrometry method for characterization of monoaromatic nitro-compounds in atmospheric particulate matter." *Journal of Chromatography A* 1268: 35-43.
- 97) Koch, B. P. and Dittmar, T. (2006). "From mass to structure: An aromaticity index for high-resolution mass data of natural organic matter." *Rapid Communications in Mass Spectrometry* 20(5): 926-932.
- 98) Koch, B. P. and Dittmar, T. (2016). "From mass to structure: An aromaticity index for high-resolution mass data of natural organic matter (vol 20, pg 926, 2006)." *Rapid Communications in Mass Spectrometry* 30(1): 250-250.
- 99) Konermann, L., Ahadi, E., Rodriguez, A. D. and Vahidi, S. (2013). "Unraveling the mechanism of electrospray ionization." *Analytical Chemistry* 85(1): 2-9.
- 100) Koppmann, R., Czapiewski, K. v. and Reid, J. (2005). "A review of biomass burning emissions, part i: Gaseous emissions of carbon monoxide, methane, volatile organic compounds, and nitrogen containing compounds." *Atmospheric chemistry and physics discussions* 5(5): 10455-10516.
- 101) Kourtchev, I., O'Connor, I. P., Giorio, C., Fuller, S. J., Kristensen, K., Maenhaut, W., Wenger, J. C., Sodeau, J. R., Glasius, M. and Kalberer, M. (2014). "Effects of anthropogenic emissions on the molecular composition of urban organic aerosols: An ultrahigh resolution mass spectrometry study." *Atmospheric Environment* 89: 525-532.
- 102) Kozhinov, A. N., Zhurov, K. O. and 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." *Analytical Chemistry* 85(13): 6437-6445.
- 103) Kroll, J. H., Donahue, N. M., Jimenez, J. L., Kessler, S. H., Canagaratna, M. R., Wilson, K. R., Altieri, K. E., Mazzoleni, L. R., Wozniak, A. S., Bluhm, H., Mysak, E. R., Smith, J. D., Kolb, C. E. and Worsnop, D. R. (2011). "Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol." *Nature Chemistry* 3(2): 133-139.
- 104) Kroll, J. H. and Seinfeld, J. H. (2008). "Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere." *Atmospheric Environment* 42(16): 3593-3624.
- 105) Kujawinski, E. B. and Behn, M. D. (2006). "Automated analysis of electrospray ionization fourier transform ion cyclotron resonance mass spectra of natural organic matter." *Analytical Chemistry* 78(13): 4363-4373.

- 106) Larsen, B. R., Gilardoni, S., Stenstrom, K., Niedzialek, J., Jimenez, J. and Belis, C. A. (2012). "Sources for pm air pollution in the po plain, italy: Ii. Probabilistic uncertainty characterization and sensitivity analysis of secondary and primary sources." *Atmospheric Environment* 50: 203-213.
- 107) Laskin, A., Gilles, M. K., Knopf, D. A., Wang, B. B. and China, S. (2016). "Progress in the analysis of complex atmospheric particles." *Annual Review of Analytical Chemistry, Vol 9* 9: 117-143.
- 108) Laskin, A., Laskin, J. and Nizkorodov, S. A. (2015). "Chemistry of atmospheric brown carbon." *Chemical Reviews* 115(10): 4335-4382.
- 109) Laskin, J., Eckert, P. A., Roach, P. J., Heath, B. S., Nizkorodov, S. A. and Laskin, A. (2012). "Chemical analysis of complex organic mixtures using reactive nanospray desorption electrospray ionization mass spectrometry." *Analytical Chemistry* 84(16): 7179-7187.
- 110) LeClair, J. P., Collett, J. L. and Mazzoleni, L. R. (2012). "Fragmentation analysis of water-soluble atmospheric organic matter using ultrahigh-resolution ft-icr mass spectrometry." *Environmental Science & Technology* 46(8): 4312-4322.
- 111) Lee, A. K. Y., Hayden, K. L., Herckes, P., Leaitch, W. R., Liggio, J., Macdonald, A. M. and Abbatt, J. P. D. (2012). "Characterization of aerosol and cloud water at a mountain site during wacs 2010: Secondary organic aerosol formation through oxidative cloud processing." *Atmospheric Chemistry and Physics* 12(15): 7103-7116.
- 112) Lee, A. K. Y., Zhao, R., Li, R., Liggio, J., Li, S. M. and Abbatt, J. P. D. (2013). "Formation of light absorbing organo-nitrogen species from evaporation of droplets containing glyoxal and ammonium sulfate." *Environmental Science & Technology* 47(22): 12819-12826.
- 113) Lee, C. G., Yuan, C. S., Chang, J. C. and Yuan, C. (2005). "Effects of aerosol species on atmospheric visibility in kaohsiung city, taiwan." *Journal of the Air & Waste Management Association* 55(7): 1031-1041.
- 114) Li, C. L., He, Q. F., Schade, J., Passig, J., Zimmermann, R., Meidan, D., Laskin, A. and Rudich, Y. (2019). "Dynamic changes in optical and chemical properties of tar ball aerosols by atmospheric photochemical aging." *Atmospheric Chemistry and Physics* 19(1): 139-163.
- 115) Li, Y., Poschl, U. and Shiraiwa, M. (2016). "Molecular corridors and parameterizations of volatility in the chemical evolution of organic aerosols." *Atmospheric Chemistry and Physics* 16(5): 3327-3344.

- 116) Lima, A. L. C., Farrington, J. W. and Reddy, C. M. (2005). "Combustion-derived polycyclic aromatic hydrocarbons in the environment - a review." *Environmental Forensics* 6(2): 109-131.
- 117) Lin, P., Bluvshstein, N., Rudich, Y., Nizkorodov, S. A., Laskin, J. and Laskin, A. (2017). "Molecular chemistry of atmospheric brown carbon inferred from a nationwide biomass burning event." *Environmental Science & Technology* 51(20): 11561-11570.
- 118) Lin, P., Fleming, L. T., Nizkorodov, S. A., Laskin, J. and Laskin, A. (2018). "Comprehensive molecular characterization of atmospheric brown carbon by high resolution mass spectrometry with electrospray and atmospheric pressure photoionization." *Analytical Chemistry* 90(21): 12493-12502.
- 119) Lin, P., Rincon, A. G., Kalberer, M. and Yu, J. Z. (2012). "Elemental composition of hulis in the pearl river delta region, china: Results inferred from positive and negative electrospray high resolution mass spectrometric data." *Environmental Science & Technology* 46(14): 7454-7462.
- 120) Lin, P. and Yu, J. Z. (2011). "Generation of reactive oxygen species mediated by humic-like substances in atmospheric aerosols." *Environmental Science & Technology* 45(24): 10362-10368.
- 121) Lohmann, U. (2006). "Aerosol effects on clouds and climate." *Space Science Reviews* 125(1-4): 129-137.
- 122) Malloy, Q., Qi, L., Warren, B., Cocker Iii, D., Erupe, M. and Silva, P. (2009). "Secondary organic aerosol formation from primary aliphatic amines with no 3 radical." *Atmospheric Chemistry and Physics* 9(6): 2051-2060.
- 123) Marshall, A. G. (2000). "Milestones in fourier transform ion cyclotron resonance mass spectrometry technique development." *International Journal of Mass Spectrometry* 200(1-3): 331-356.
- 124) Mazzoleni, L. R., Ehrmann, B. M., Shen, X. H., Marshall, A. G. and 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.
- 125) Mazzoleni, L. R., Saranjampour, P., Dalbec, M. M., Samburova, V., Hallar, A. G., Zielinska, B., Lowenthal, D. H. and Kohl, S. (2012). "Identification of water-soluble organic carbon in non-urban aerosols using ultrahigh-resolution ft-icr mass spectrometry: Organic anions." *Environmental Chemistry* 9(3): 285-297.

- 126) Mazzoleni, L. R., Zielinska, B. and Moosmuller, H. (2007). "Emissions of levoglucosan, methoxy phenols, and organic acids from prescribed burns, laboratory combustion of wildland fuels, and residential wood combustion." *Environmental Science & Technology* 41(7): 2115-2122.
- 127) McNeill, V. F. (2015). "Aqueous organic chemistry in the atmosphere: Sources and chemical processing of organic aerosols." *Environmental Science & Technology* 49(3): 1237-1244.
- 128) McNeill, V. F., Woo, J. L., Kim, D. D., Schwier, A. N., Wannell, N. J., Sumner, A. J. and Barakat, J. M. (2012). "Aqueous-phase secondary organic aerosol and organosulfate formation in atmospheric aerosols: A modeling study." *Environmental Science & Technology* 46(15): 8075-8081.
- 129) Moosmuller, H., Chakrabarty, R. K. and Arnott, W. P. (2009). "Aerosol light absorption and its measurement: A review." *Journal of Quantitative Spectroscopy & Radiative Transfer* 110(11): 844-878.
- 130) Moses, E. E. and Akinyemi, M. (2017). A short review on the effects of aerosols on visibility impairment. Journal of Physics: Conference Series, IOP Publishing.
- 131) Nguyen, T. B., Laskin, A., Laskin, J. and Nizkorodov, S. A. (2013). "Brown carbon formation from ketoaldehydes of biogenic monoterpenes." *Faraday Discussions* 165: 473-494.
- 132) Nguyen, T. B., Lee, P. B., Updyke, K. M., Bones, D. L., Laskin, J., Laskin, A. and Nizkorodov, S. A. (2012). "Formation of nitrogen- and sulfur-containing light-absorbing compounds accelerated by evaporation of water from secondary organic aerosols." *Journal of Geophysical Research-Atmospheres* 117.
- 133) Nizkorodov, S. A., Laskin, J. and Laskin, A. (2011). "Molecular chemistry of organic aerosols through the application of high resolution mass spectrometry." *Physical Chemistry Chemical Physics* 13(9): 3612-3629.
- 134) Noziere, B., Ekstrom, S., Alsberg, T. and Holmstrom, S. (2010). "Radical-initiated formation of organosulfates and surfactants in atmospheric aerosols." *Geophysical Research Letters* 37: L05806.
- 135) Noziere, B., Kaberer, M., Claeys, M., Allan, J., D'Anna, B., Decesari, S., Finessi, E., Glasius, M., Grgic, I., Hamilton, J. F., Hoffmann, T., Iinuma, Y., Jaoui, M., Kahno, A., Kampf, C. J., Kourtev, I., Maenhaut, W., Marsden, N., Saarikoski, S., Schnelle-Kreis, J., Surratt, J. D., Szidat, S., Szmigielski, R. and Wisthaler, A. (2015). "The molecular identification of organic compounds in the atmosphere: State of the art and challenges." *Chemical Reviews* 115(10): 3919-3983.
- 136) NWCC (2018). Northwest annual fire report 2017.

- 137) Paglione, M., Saarikoski, S., Carbone, S., Hillamo, R., Facchini, M. C., Finessi, E., Giulianelli, L., Carbone, C., Fuzzi, S., Moretti, F., Tagliavini, E., Swietlicki, E., Stenstrom, K. E., Prevot, A. S. H., Massoli, P., Canaragatna, M., Worsnop, D. and Decesari, S. (2014). "Primary and secondary biomass burning aerosols determined by proton nuclear magnetic resonance (h-1-nmr) spectroscopy during the 2008 eucaari campaign in the po valley (italy)." *Atmospheric Chemistry and Physics* 14(10): 5089-5110.
- 138) Pavia, D., Lampman, G., Kriz, G., Vyvyan, J. (2009). Introduction to spectroscopy, Brooks/Cole, Cengage Learning.
- 139) Phillips, S. M. and Smith, G. D. (2014). "Light absorption by charge transfer complexes in brown carbon aerosols." *Environmental Science & Technology Letters* 1(10): 382-386.
- 140) Phillips, S. M. and Smith, G. D. (2015). "Further evidence for charge transfer complexes in brown carbon aerosols from excitation-emission matrix fluorescence spectroscopy." *Journal of Physical Chemistry A* 119(19): 4545-4551.
- 141) Pietrogrande, M. C., Bacco, D., Visentin, M., Ferrari, S. and Casali, P. (2014). "Polar organic marker compounds in atmospheric aerosol in the po valley during the supersito campaigns - part 2: Seasonal variations of sugars." *Atmospheric Environment* 97: 215-225.
- 142) Pietrogrande, M. C., Bacco, D., Visentin, M., Ferrari, S. and Poluzzi, V. (2014). "Polar organic marker compounds in atmospheric aerosol in the po valley during the supersito campaigns - part 1: Low molecular weight carboxylic acids in cold seasons." *Atmospheric Environment* 86: 164-175.
- 143) Pietrogrande, M. C., Sacco, D., Ferrari, S., Kaipainen, J., Ricciardelli, I., Riekkola, M. L., Trentini, A. and Visentin, M. (2015). "Characterization of atmospheric aerosols in the po valley during the supersito campaigns - part 3: Contribution of wood combustion to wintertime atmospheric aerosols in emilia romagna region (northern italy)." *Atmospheric Environment* 122: 291-305.
- 144) Poluzzi, V., Trentini, A., Scotto, F., Ricciardelli, I., Ferrari, S., Maccone, C., Bacco, D., Zigola, C., Bonafè, G. and Ugolini, P. (2015). "Preliminary results of the project "supersito" concerning the atmospheric aerosol composition in emilia-romagna region, italy: Pm source apportionment and aerosol size distribution." *WIT Transactions on The Built Environment* 168: 689-698.
- 145) Poschl, U. (2005). "Atmospheric aerosols: Composition, transformation, climate and health effects." *Angewandte Chemie-International Edition* 44(46): 7520-7540.

- 146) Poschl, U. and Shiraiwa, M. (2015). "Multiphase chemistry at the atmosphere-biosphere interface influencing climate and public health in the anthropocene." *Chemical Reviews* 115(10): 4440-4475.
- 147) Posfai, M., Gelencser, A., Simonics, R., Arato, K., Li, J., Hobbs, P. V. and Buseck, P. R. (2004). "Atmospheric tar balls: Particles from biomass and biofuel burning." *Journal of Geophysical Research-Atmospheres* 109(D6).
- 148) Pratt, K. A., Fiddler, M. N., Shepson, P. B., Carlton, A. G. and Surratt, J. D. (2013). "Organosulfates in cloud water above the ozarks' isoprene source region." *Atmospheric Environment* 77: 231-238.
- 149) Putman, A. L., Offenberg, J. H., Fisseha, R., Kundu, S., Rahn, T. A. and Mazzoleni, L. R. (2012). "Ultrahigh-resolution ft-icr mass spectrometry characterization of alpha-pinene ozonolysis soa." *Atmospheric Environment* 46: 164-172.
- 150) Reid, J. P., Bertram, A. K., Topping, D. O., Laskin, A., Martin, S. T., Petters, M. D., Pope, F. D. and Rovelli, G. (2018). "The viscosity of atmospherically relevant organic particles." *Nature Communications* 9.
- 151) Reid, J. S., Koppmann, R., Eck, T. F. and Eleuterio, D. P. (2005). "A review of biomass burning emissions part ii: Intensive physical properties of biomass burning particles." *Atmospheric Chemistry and Physics* 5: 799-825.
- 152) Riccobono, F., Schobesberger, S., Scott, C. E., Dommen, J., Ortega, I. K., Rondo, L., Almeida, J., Amorim, A., Bianchi, F., Breitenlechner, M., David, A., Downard, A., Dunne, E. M., Duplissy, J., Ehrhart, S., Flagan, R. C., Franchin, A., Hansel, A., Junninen, H., Kajos, M., Keskinen, H., Kupc, A., Kurten, A., Kvashin, A. N., Laaksonen, A., Lehtipalo, K., Makhmutov, V., Mathot, S., Nieminen, T., Onnela, A., Petaja, T., Praplan, A. P., Santos, F. D., Schallhart, S., Seinfeld, J. H., Sipila, M., Spracklen, D. V., Stozhkov, Y., Stratmann, F., Tome, A., Tsagkogeorgas, G., Vaattovaara, P., Viisanen, Y., Vrtala, A., Wagner, P. E., Weingartner, E., Wex, H., Wimmer, D., Carslaw, K. S., Curtius, J., Donahue, N. M., Kirkby, J., Kulmala, M., Worsnop, D. R. and Baltensperger, U. (2014). "Oxidation products of biogenic emissions contribute to nucleation of atmospheric particles." *Science* 344(6185): 717-721.
- 153) Rinaldi, M., Emblico, L., Decesari, S., Fuzzi, S., Facchini, M. C. and Librando, V. (2007). "Chemical characterization and source apportionment of size-segregated aerosol collected at an urban site in sicily." *Water Air and Soil Pollution* 185(1-4): 311-321.
- 154) Riva, M., Budisulistiorini, S. H., Zhang, Z. F., Gold, A., Thornton, J. A., Turpin, B. J. and Surratt, J. D. (2017). "Multiphase reactivity of gaseous hydroperoxide oligomers produced from isoprene ozonolysis in the presence of acidified aerosols." *Atmospheric Environment* 152: 314-322.

- 155) Rosenfeld, D. (2006). "Aerosol-cloud interactions control of earth radiation and latent heat release budgets." *Space Science Reviews* 125(1-4): 149-157.
- 156) Saarikoski, S., Carbone, S., Decesari, S., Giulianelli, L., Angelini, F., Canagaratna, M., Ng, N. L., Trimborn, A., Facchini, M. C., Fuzzi, S., Hillamo, R. and Worsnop, D. (2012). "Chemical characterization of springtime submicrometer aerosol in po valley, italy." *Atmospheric Chemistry and Physics* 12(18): 8401-8421.
- 157) Salby, M. L. (2012). Physics of the atmosphere and climate, Cambridge University Press.
- 158) Saleh, R., Cheng, Z. Z. and Atwi, K. (2018). "The brown-black continuum of light-absorbing combustion aerosols." *Environmental Science & Technology Letters* 5(8): 508-513.
- 159) Samburova, V., Connolly, J., Gyawali, M., Yatavelli, R. L. N., Watts, A. C., Chakrabarty, R. K., Zielinska, B., Moosmuller, H. and Khlystov, A. (2016). "Polycyclic aromatic hydrocarbons in biomass-burning emissions and their contribution to light absorption and aerosol toxicity." *Science of the Total Environment* 568: 391-401.
- 160) Samburova, V., Zielinska, B. and Khlystov, A. (2017). "Do 16 polycyclic aromatic hydrocarbons represent pah air toxicity?" *Toxics* 5(3).
- 161) Sareen, N., Waxman, E. M., Turpin, B. J., Volkamer, R. and Carlton, A. G. (2017). "Potential of aerosol liquid water to facilitate organic aerosol formation: Assessing knowledge gaps about precursors and partitioning." *Environmental Science & Technology* 51(6): 3327-3335.
- 162) Savory, J. J., Kaiser, N. K., McKenna, A. M., Xian, F., Blakney, G. T., Rodgers, R. P., Hendrickson, C. L. and Marshall, A. G. (2011). "Parts-per-billion fourier transform ion cyclotron resonance mass measurement accuracy with a "walking" calibration equation." *Analytical Chemistry* 83(5): 1732-1736.
- 163) Schindelka, J., Iinuma, Y., Hoffmann, D. and Herrmann, H. (2013). "Sulfate radical-initiated formation of isoprene-derived organosulfates in atmospheric aerosols." *Faraday Discussions* 165: 237-259.
- 164) Schmitt-Kopplin, P., Gelencser, A., Dabek-Zlotorzynska, E., Kiss, G., Hertkorn, N., Harir, M., Hong, Y. and Gebefugi, I. (2010). "Analysis of the unresolved organic fraction in atmospheric aerosols with ultrahigh-resolution mass spectrometry and nuclear magnetic resonance spectroscopy: Organosulfates as photochemical smog constituents." *Analytical Chemistry* 82(19): 8017-8026.

- 165) Schum, S. K., Mazzoleni, L. R. and Brown, L. E. (2018). "Mfassignr: An r package for data preparation and molecular formula assignment."
- 166) Schum, S. K., Zhang, B., Dzepina, K., Fialho, P., Mazzoleni, C. and Mazzoleni, L. R. (2018). "Molecular and physical characteristics of aerosol at a remote free troposphere site: Implications for atmospheric aging." *Atmospheric Chemistry and Physics* 18(19): 14017-14036.
- 167) Schwartz, J., Laden, F. and Zanobetti, A. (2002). "The concentration-response relation between pm2.5 and daily deaths." *Environmental Health Perspectives* 110(10): 1025-1029.
- 168) Sedlacek, A. J., Buseck, P. R., Adachi, K., Onasch, T. B., Springston, S. R. and Kleinman, L. (2018). "Formation and evolution of tar balls from northwestern us wildfires." *Atmospheric Chemistry and Physics* 18(15): 11289-11301.
- 169) See, S. W., Wang, Y. H. and Balasubramanian, R. (2007). "Contrasting reactive oxygen species and transition metal concentrations in combustion aerosols." *Environmental Research* 103(3): 317-324.
- 170) Seinfeld, J. H. and Pandis, S. N. (2016). Atmospheric chemistry and physics: From air pollution to climate change, John Wiley & Sons.
- 171) Sharma, M. and Maloo, S. (2005). "Assessment of ambient air pm10 and pm2.5 and characterization of pm10 in the city of kanpur, india." *Atmospheric Environment* 39(33): 6015-6026.
- 172) Shimizu, M. and Yano, E. (1986). "Mutagenicity of mono-nitrobenzene derivatives in the ames test and rec assay." *Mutation Research* 170(1-2): 11-22.
- 173) Shinozuka, Y., Clarke, A. D., Howell, S. G., Kapustin, V. N., McNaughton, C. S., Zhou, J. C. and Anderson, B. E. (2007). "Aircraft profiles of aerosol microphysics and optical properties over north america: Aerosol optical depth and its association with pm2.5 and water uptake." *Journal of Geophysical Research-Atmospheres* 112(D12).
- 174) Shiraiwa, M., Kondo, Y., Iwamoto, T. and Kita, K. (2010). "Amplification of light absorption of black carbon by organic coating." *Aerosol Science and Technology* 44(1): 46-54.
- 175) Shiraiwa, M., Li, Y., Tsimpidi, A. P., Karydis, V. A., Berkemeier, T., Pandis, S. N., Lelieveld, J., Koop, T. and Poschl, U. (2017). "Global distribution of particle phase state in atmospheric secondary organic aerosols." *Nature Communications* 8.

- 176) Shrivastava, M., Cappa, C. D., Fan, J. W., Goldstein, A. H., Guenther, A. B., Jimenez, J. L., Kuang, C., Laskin, A., Martin, S. T., Ng, N. L., Petaja, T., Pierce, J. R., Rasch, P. J., Roldin, P., Seinfeld, J. H., Shilling, J., Smith, J. N., Thornton, J. A., Volkamer, R., Wang, J., Worsnop, D. R., Zaveri, R. A., Zelenyuk, A. and Zhang, Q. (2017). "Recent advances in understanding secondary organic aerosol: Implications for global climate forcing." *Reviews of Geophysics* 55(2): 509-559.
- 177) Shrivastava, M., Lou, S., Zelenyuk, A., Easter, R. C., Corley, R. A., Thrall, B. D., Rasch, P. J., Fast, J. D., Simonich, S. L. M. and Shen, H. (2017). "Global long-range transport and lung cancer risk from polycyclic aromatic hydrocarbons shielded by coatings of organic aerosol." *Proceedings of the National Academy of Sciences* 114(6): 1246-1251.
- 178) Simoneit, B. R. T. (2002). "Biomass burning - a review of organic tracers for smoke from incomplete combustion." *Applied Geochemistry* 17(3): 129-162.
- 179) Staudt, S., Kundu, S., Lehmler, H. J., He, X. R., Cui, T. Q., Lin, Y. H., Kristensen, K., Glasius, M., Zhang, X. L., Weber, R. J., Surratt, J. D. and Stone, E. A. (2014). "Aromatic organosulfates in atmospheric aerosols: Synthesis, characterization, and abundance." *Atmospheric Environment* 94: 366-373.
- 180) Stenson, A. C., Marshall, A. G. and 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." *Analytical Chemistry* 75(6): 1275-1284.
- 181) Sun, H. L., Biedermann, L. and Bond, T. C. (2007). "Color of brown carbon: A model for ultraviolet and visible light absorption by organic carbon aerosol." *Geophysical Research Letters* 34(17).
- 182) Surratt, J. D., Gomez-Gonzalez, Y., Chan, A. W. H., Vermeylen, R., Shahgholi, M., Kleindienst, T. E., Edney, E. O., Offenberg, J. H., Lewandowski, M., Jaoui, M., Maenhaut, W., Claeys, M., Flagan, R. C. and Seinfeld, J. H. (2008). "Organosulfate formation in biogenic secondary organic aerosol." *Journal of Physical Chemistry A* 112(36): 8345-8378.
- 183) Suzuki, Y., Kawakami, M. and Akasaka, K. (2001). "H-1 nmr application for characterizing water-soluble organic compounds in urban atmospheric particles." *Environmental Science & Technology* 35(13): 2656-2664.
- 184) Tfaily, M. M., Chu, R. K., Tolic, N., Roscioli, K. M., Anderton, C. R., Pasa-Tolic, L., Robinson, E. W. and Hess, N. J. (2015). "Advanced solvent based methods for molecular characterization of soil organic matter by high-resolution mass spectrometry." *Analytical Chemistry* 87(10): 5206-5215.

- 185) Tivanski, A. V., Hopkins, R. J., Tyliczszak, T. and Gilles, M. K. (2007). "Oxygenated interface on biomass burn tar balls determined by single particle scanning transmission x-ray microscopy." *Journal of Physical Chemistry A* 111(25): 5448-5458.
- 186) Tong, H. J., Kourtchev, I., Pant, P., Keyte, I. J., O'Connor, I. P., Wenger, J. C., Pope, F. D., Harrison, R. M. and Kalberer, M. (2016). "Molecular composition of organic aerosols at urban background and road tunnel sites using ultra-high resolution mass spectrometry." *Faraday Discussions* 189: 51-68.
- 187) Trenberth, K. E., Fasullo, J. T. and Kiehl, J. (2009). "Earth's global energy budget." *Bulletin of the American Meteorological Society* 90(3): 311-323.
- 188) Tu, P. J., Hall, W. A. and Johnston, M. V. (2016). "Characterization of highly oxidized molecules in fresh and aged biogenic secondary organic aerosol." *Analytical Chemistry* 88(8): 4495-4501.
- 189) Twomey, S. (1977). "Influence of pollution on shortwave albedo of clouds." *Journal of the Atmospheric Sciences* 34(7): 1149-1152.
- 190) Vaden, T. D., Imre, D., Beranek, J., Shrivastava, M. and Zelenyuk, A. (2011). "Evaporation kinetics and phase of laboratory and ambient secondary organic aerosol." *Proceedings of the National Academy of Sciences of the United States of America* 108(6): 2190-2195.
- 191) Valavanidis, A., Fiotakis, K. and Vlachogianni, T. (2010). The role of stable free radicals, metals and pahs of airborne particulate matter in mechanisms of oxidative stress and carcinogenicity. *Urban airborne particulate matter*, Springer: 411-426.
- 192) Verma, V., Fang, T., Xu, L., Peltier, R. E., Russell, A. G., Ng, N. L. and Weber, R. J. (2015). "Organic aerosols associated with the generation of reactive oxygen species (ros) by water-soluble pm2.5." *Environmental Science & Technology* 49(7): 4646-4656.
- 193) Wang, J., Shilling, J. E., Liu, J. M., Zelenyuk, A., Bell, D. M., Petters, M. D., Thalman, R., Mei, F., Zaveri, R. A. and Zheng, G. J. (2019). "Cloud droplet activation of secondary organic aerosol is mainly controlled by molecular weight, not water solubility." *Atmospheric Chemistry and Physics* 19(2): 941-954.
- 194) Wang, L., Li, Z. Q., Tian, Q. J., Ma, Y., Zhang, F. X., Zhang, Y., Li, D. H., Li, K. T. and Li, L. (2013). "Estimate of aerosol absorbing components of black carbon, brown carbon, and dust from ground-based remote sensing data of sun-sky radiometers." *Journal of Geophysical Research-Atmospheres* 118(12): 6534-6543.

- 195) Wang, X., Doherty, S. J. and Huang, J. P. (2013). "Black carbon and other light-absorbing impurities in snow across northern china." *Journal of Geophysical Research-Atmospheres* 118(3): 1471-1492.
- 196) Wang, X. K., Hayeck, N., Brüggemann, M., Yao, L., Chen, H. F., Zhang, C., Emmelin, C., Chen, J. M., George, C. and Wang, L. (2017). "Chemical characteristics of organic aerosols in shanghai: A study by ultrahigh-performance liquid chromatography coupled with orbitrap mass spectrometry." *Journal of Geophysical Research-Atmospheres* 122(21): 11703-11722.
- 197) Warner, J. X., Dickerson, R. R., Wei, Z., Strow, L. L., Wang, Y. and Liang, Q. (2017). "Increased atmospheric ammonia over the world's major agricultural areas detected from space." *Geophysical Research Letters* 44(6): 2875-2884.
- 198) Weimer, S., Drewnick, F., Högrefe, O., Schwab, J. J., Rhoads, K., Orsini, D., Canagaratna, M., Worsnop, D. R. and Demerjian, K. L. (2006). "Size-selective nonrefractory ambient aerosol measurements during the particulate matter technology assessment and characterization study - new york 2004 winter intensive in new york city." *Journal of Geophysical Research-Atmospheres* 111(D18).
- 199) Westerling, A. L., Hidalgo, H. G., Cayan, D. R. and Swetnam, T. W. (2006). "Warming and earlier spring increase western us forest wildfire activity." *Science* 313(5789): 940-943.
- 200) WHO (2016). "Ambient air pollution: A global assessment of exposure and burden of disease."
- 201) Wielgosiński, G. (2012). Pollutant formation in combustion processes. *Advances in chemical engineering*, IntechOpen.
- 202) Willoughby, A. S., Wozniak, A. S. and Hatcher, P. G. (2016). "Detailed source-specific molecular composition of ambient aerosol organic matter using ultrahigh resolution mass spectrometry and h-1 nmr." *Atmosphere* 7(6): 79.
- 203) Winterhalter, R., Van Dingenen, R., Larsen, B. R., Jensen, N. R. and Hjorth, J. (2003). "Lc-ms analysis of aerosol particles from the oxidation of α -pinene by ozone and oh-radicals." *Atmos. Chem. Phys. Discuss.* 2003: 1-39.
- 204) Wozniak, A. S., Willoughby, A. S., Gurganus, S. C. and Hatcher, P. G. (2014). "Distinguishing molecular characteristics of aerosol water soluble organic matter from the 2011 trans-north atlantic us geochemical cruise." *Atmospheric Chemistry and Physics* 14(16): 8419-8434.

- 205) Xu, X., Chen, J., Zhu, C., Li, J., Sui, X., Liu, L. and Sun, J. (2017). "Fog composition along the yangtze river basin: Detecting emission sources of pollutants in fog water." *Journal of Environmental Sciences*: 11.
- 206) Xue, W. L. and Warshawsky, D. (2005). "Metabolic activation of polycyclic and heterocyclic aromatic hydrocarbons and DNA damage: A review." *Toxicology and Applied Pharmacology* 206(1): 73-93.
- 207) Ye, B. M., Ji, X. L., Yang, H. Z., Yao, X. H., Chan, C. K., Cadle, S. H., Chan, T. and Mulawa, P. A. (2003). "Concentration and chemical composition of pm2.5 in shanghai for a 1-year period." *Atmospheric Environment* 37(4): 499-510.
- 208) Yin, Y., Jia, H. X., Sun, Y. Y., Yu, H. X., Wang, X. R., Wu, J. C. and Xue, Y. Q. (2007). "Bioaccumulation and ros generation in liver of carassius auratus, exposed to phenanthrene." *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology* 145(2): 288-293.
- 209) Yu, L., Smith, J., Laskin, A., Anastasio, C., Laskin, J. and Zhang, Q. (2014). "Chemical characterization of soa formed from aqueous-phase reactions of phenols with the triplet excited state of carbonyl and hydroxyl radical." *Atmospheric Chemistry and Physics* 14(24): 13801-13816.
- 210) Yu, L., Smith, J., Laskin, A., George, K. M., Anastasio, C., Laskin, J., Dillner, A. M. and Zhang, Q. (2016). "Molecular transformations of phenolic soa during photochemical aging in the aqueous phase: Competition among oligomerization, functionalization, and fragmentation." *Atmospheric Chemistry and Physics* 16(7): 4511-4527.
- 211) Yu, P., Toon, O. B., Bardeen, C. G., Zhu, Y., Rosenlof, K. H., Portmann, R. W., Thornberry, T. D., Gao, R.-S., Davis, S. M. and Wolf, E. T. J. S. (2019). "Black carbon loft wildfire smoke high into the stratosphere to form a persistent plume." *365(6453)*: 587-590.
- 212) Zark, M., Christoffers, J. and 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.
- 213) Zelenyuk, A., Imre, D., Wilson, J., Zhang, Z. Y., Wang, J. and Mueller, K. (2015). "Airborne single particle mass spectrometers (splat ii & minisplat) and new software for data visualization and analysis in a geo-spatial context." *Journal of the American Society for Mass Spectrometry* 26(2): 257-270.
- 214) Zelenyuk, A., Imre, D. G., Wilson, J., Bell, D. M., Suski, K. J., Shrivastava, M., Beranek, J., Alexander, M. L., Kramer, A. L. and Simonich, S. L. M. (2017). "The effect of gas-phase polycyclic aromatic hydrocarbons on the formation and

properties of biogenic secondary organic aerosol particles." *Faraday Discussions* 200: 143-164.

- 215) Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Allan, J. D., Coe, H., Ulbrich, I., Alfarra, M. R., Takami, A., Middlebrook, A. M., Sun, Y. L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P. F., Salcedo, D., Onasch, T., Jayne, J. T., Miyoshi, T., Shimojo, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R. J., Rautiainen, J., Sun, J. Y., Zhang, Y. M. and Worsnop, D. R. (2007). "Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced northern hemisphere midlatitudes." *Geophysical Research Letters* 34(13): L13801.
- 216) Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., Worsnop, D. R. and Sun, Y. L. (2011). "Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: A review." *Analytical and Bioanalytical Chemistry* 401(10): 3045-3067.
- 217) Zhang, R. (2010). "Getting to the critical nucleus of aerosol formation." *Science* 328(5984): 1366-1367.
- 218) Zhao, Y., Hallar, A. G. and Mazzoleni, L. R. (2013). "Atmospheric organic matter in clouds: Exact masses and molecular formula identification using ultrahigh-resolution ft-icr mass spectrometry." *Atmospheric Chemistry and Physics* 13(24): 12343-12362.
- 219) Zubarev, R. A. and Makarov, A. (2013). "Orbitrap mass spectrometry." *Analytical Chemistry* 85(11): 5288-5296.