Spring 2017

Alicyclic and Aromatic Carboxylic Acids in Soil Organic Matter: An Investigation of Potential Origin and Association with Plutonium Using Advanced Analytical Techniques

Nicole DiDonato
Old Dominion University

Recommended Citation

ALICYCLIC AND AROMATIC CARBOXYLIC ACIDS IN SOIL ORGANIC MATTER:

AN INVESTIGATION OF POTENTIAL ORIGIN AND ASSOCIATION WITH

PLUTONIUM USING ADVANCED ANALYTICAL TECHNIQUES

by

Nicole DiDonato
Bachelor of Environmental Engineering, May 2000, University of Delaware

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

CHEMISTRY

OLD DOMINION UNIVERSITY
May 2017

Approved by:

Patrick G. Hatcher (Director)

Jingdong Mao (Member)

John Donat (Member)

Gary Schafran (Member)

ABSTRACT

ALICYCLIC AND AROMATIC CARBOXYLIC ACIDS IN SOIL ORGANIC MATTER: AN INVESTIGATION OF POTENTIAL ORIGIN AND ASSOCIATION WITH PLUTONIUM USING ADVANCED ANALYTICAL TECHNIQUES

Nicole DiDonato
Old Dominion University, 2017
Director: Dr. Patrick G. Hatcher

Carboxylic acids are a defining component of soil organic matter, responsible for many of the physical and chemical properties, including metal-organic matter interactions, which govern its role as an important constituent of soils. However, there is a shortage of detailed molecular level information regarding orientation and structural arrangement of carboxylic acids within soil organic matter. This dissertation utilizes electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICRMS) as well as solid-state and multi-dimensional nuclear magnetic resonance (NMR) to investigate the molecular formula composition within several organic matter sources and the primary structures that feature carboxylic acids. Soil organic matter is evaluated in two forms: as the alkali-soluble, acid-insoluble portion of organic matter (humic acids) from a collection of sources, as well as the alkali soluble soil organic material associated with high Fe and Pu metal concentrations at a contaminated munitions facility. Two predominant carboxyl-containing molecular assemblages are found to be common in a wide variety of soil humic acids. Along with lignin-like assemblages, these are carboxyl-containing aliphatic molecules (CCAM) and condensed aromatic molecules. The proportion of these groups relative to lignin-like compounds within samples and the percent of total carboxylic acid molecular formulas among samples are found to

increase with increasing humification of the soil. Since CCAM and condensed aromatic molecules have previously been shown to be generated from oxidization of lignin, this represents renewed evidence for lignin as a major source of organic matter in soils. Lignin ring-opening and radical re-polymerization reactions have been proposed to form alicyclic CCAM and condensed aromatic molecules. Detailed evaluation of the aliphatic molecules using multi-dimensional NMR confirms the presence of ring structures, replete with carboxylic acids, heteroatom substitutions in the form of alcohols and ethers, as well as a variety of methyl group substituents. Additionally, condensed aromatic carboxylic acid molecular formulas, primarily those containing nitrogen, were found composed in organic matter with elevated metal ions Pu and Fe. Carboxylic acid oxygens in combination with nitrogen in aromatic structures are suspected to be partially responsible for the high metal affinity. Nitrogen-containing hydroxamate groups were also investigated for their potential to be incorporated into stable organic matter by testing the reaction between an amine-containing hydroxamate siderophore and the biopolymer cutin. While products of this reaction could not be confirmed, carboxylic acid functional groups are identified in this thesis as key molecular components contributing to Pu and Fe metal-binding attributes of organic matter, and potentially formed during the production of condensed aromatic and alicyclic compounds as a result of radical oxidation reactions of lignin.

To my family, for their infinite love and support.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Soil organic matter (SOM) constitutes a critical pool of organic carbon, residing on the earth's surface at the interface between air, water, rock and living organisms. Understanding its role at the boundary between these entities is crucial not only for the success of agriculture in supporting the exponential growth of the world's population, but also for understanding global warming and the micro and macro-cycles of nutrients and pollutants. Soil organic matter properties such as structure and composition, along with ecosystem and environmental factors are key to understanding these processes. However, of the 1600 Gtons of carbon that resides in soils, the largest pool of dynamic organic carbon on Earth, nearly two thirds of this has been estimated to be composed primarily of a molecularly uncharacterized fraction (Bianchi, 2011; Hedges et al., 1997). The remainder is made up of known biomolecules such as cellulose (250 G ton), lignin (175 G ton), polysaccharides (150 G ton), and other minor components such as proteins, lipids, and cutin (5-10 G ton) which are presumed to be the primary sources leading to less-recognizable forms upon degradation (Bianchi, 2011). Humic substances, commonly subdivided as humic acids, fulvic acids, dissolved organic matter, and insoluble humin collectively account for the bulk of this uncharacterized fraction. Humic acids, fulvic acids and humin are operationally defined based on their solubility or insolubility in acid and alkali solutions. Humic acids are soluble in base, but precipitate at low pH while fulvic acids are soluble in acid and under alkali conditions. Humin is insoluble at all pH ranges. These operational methods of sorting organic matter for analysis are an important alternative to traditional attempts which rely heavily on analytical techniques that provide average functional

---

This dissertation is formatted based on the journal *Organic Geochemistry*.

group compositions or alter the organic matter by physical separations and chemical and/or thermal means to render the fractions more amenable to molecular-level analysis. An additional advantage of using alkali is that acid functional groups such as carboxylic acids that are characteristic of humic substances are ionized over a wide pH range (<2-11, in general for fulvic acids) and those that are otherwise insoluble but can ionize at high pH to facilitate solubility in base (i.e. humic acids) to be extracted. Thus, alkali extracted organic matter that is subsequently acid precipitated to yield humic acids is the focal point for analysis in this thesis.

The structural configurations of humic substances are integrally linked to how they are formed from parent molecules. There is no shortage of proposed theories and mechanisms for the formation of the uncharacterized portion of organic matter. While there is little consensus among the scientific community, some of the oldest theories still considered feasible, either in whole or in combination with other proposed mechanisms, include the lignin theory in which humic and fulvic acids are formed after microbial oxidation of lignin via depolymerization, demethylation and oxidation to form catechols as well as condensation with nitrogen compounds (Stevenson, 1994; Flaig, 1964). Alternatively, the lignin decomposition or polyphenol theory, is also still considered to have some merit (Sparks, 2003; Haider and Martin, 1967).   According to this theory, microbial degradation may also include non-lignin carbon sources to form phenolic aldehydes and acids, which are further altered to polyphenols and quinones that have been proposed to incorporate nitrogen abiotically into humic materials via Michael adducts or Schiff bases (Stevenson, 1994). The sugar-amine or melanoidin pathway in which amines add to sugar aldehydes to form humic materials is also still considered possible, but not a major process (Sparks, 2003). Phenolic products have been suggested to condense and re-polymerize to form

large covalently linked molecules, or merely aggregate as a result of strong intramolecular forces (Sollins et al., 2007; Sutton and Sposito, 2005). Theories based on lignin as a precursor hold up at least in lignin dominated environments but do not explain non-lignin origins or environments where lignin biomarkers are lacking. Still other theories have been proposed for production of humic substances from tannins and terpenoids. The number of precursor molecules and combinations of reactions suggests a diversity of products that likely form a multitude of compounds, few if any of which will be identical (Stevenson, 1994). In fact there are many reports of highly varied structures in organic matter from different environments, however some remarkable similarities do exist when comparing average properties (Sparks, 2003).

Lignin is one of the most abundant biopolymers on earth, second only to cellulose and accounts for approximately 20% of the annual carbon fixed by photosynthesis, so it is not surprising it has been proposed as a major source of organic matter, as is evident from the theories described above. However these hypotheses have intermittently gained support and opposition over the course of study of humic substances. The primary arguments in refute of this theory are the lack of hydroxylated phenol biomarkers characteristic of lignin-derived structures in some environments, as well as the presence of more aliphatic compounds. However, recent work has suggested that both aliphatic and aromatic molecular assemblages may have origins from lignin, based on ESI-FTICRMS studies of lignin-derived DOM and radical oxidations of lignin (Chen et al., 2014; Waggoner et al., 2015). Chapter II of this thesis further takes advantage of the high resolving power of ESI-FTICRMS and solid state NMR to investigate the prevalence of these two molecular assemblages among soil organic matter humic acids from a variety of sources to support this hypothesis. Chapter III focuses on the structural conformation of carboxylic acid

containing alicyclic compounds in particular, in support of the findings of Chapter II using multi-dimensional NMR in combination with ESI-FTICRMS. Chapter IV is directed at investigating the condensed aromatic carboxylic acid molecular formulas which were found to be concentrated in a soil organic matter colloid containing high concentrations of the metals Pu and Fe.

Oxidation is a fundamental process in the formation of humic substances, and carboxylic acids, which are formed during the oxidation of organic matter, are abundant within humic substances and are a good indicator of oxidative processes. Carboxylic acids confer important properties to soil organic matter (SOM) such as cation exchange capacity, pH buffer capacity, metal chelation, sorptivity, solubility and reactivity that can control both nutrient and pollutant cycling in soils. The magnitude and extent of interactions, particularly for metal chelation and reactivity which are within the scope of this work, depend not only on the incidence of carboxylic functional groups in a given soil, but more so on the molecular framework and orientation in which they are incorporated into the constituent organic molecules of SOM. Unfortunately, very few naturally existing ligands have yet to be isolated and studying their molecular environment based on bulk properties is limiting, even when analytical challenges can be overcome.

Some of the most useful techniques that do not require extensive manipulations of samples (i.e. separations, extractions, derivatizations, etc.) that can alter the molecular structures and produce artifacts, are multi-dimensional NMR which provides direct structural evidence and soft ionization high resolution mass spectrometry (ESI-FTICRMS) which provides indirect structural information derived from exact molecular formulas. Heteronuclear NMR correlations between carbon and protons of single bond distance can be observed using heteronuclear single quantoum

coherence (HSQC), and heteronuclear multiple bond coherence (HMBC) can be used to detect correlations for up to three bond distances between nuclei.

Identification of naturally occurring ligand complexes within the complex matrix of environmental samples using ESI-FTICRMS has yet to be accomplished due to lack of apriori knowledge of ligand structure, composition, charge state, stoichiometry of bound metal complexes and binding strengths. Distinguishing complexes representative of natural solution conditions from artifacts of adduct formation within the ESI source also presents a challenge (McDonald et al., 2014). Known ligands such as desferoxamine are often detected as their sodium or chlorine salt adducts which are difficult to differentiate from incorrectly assigned artifacts, or "salts"(Waska et al., 2015). Smaller ligands such as citric acid also tend to form multi-nuclear complexes consisting of many stoichiometric combinations of citrate molecules and metals (Gautier-Luneau et al., 2005). However the technique is useful for identifying potential un-bound ligands, particularly those containing functional groups of interest such as carboxylic acids which many natural organic matter ligands are thought to contain (Hertkorn et al., 2006; Repeta, 2015). One way to do this is to successively 'pool' organic matter that is most closely bound or associated with metal species while simultaneously eliminating other organic matter structures that do not interact with metals. Ultimately a high concentration of metals and natural organic matter ligands are obtained. This was done for soil contaminated with plutonium and containing high amounts of iron (Xu et al., 2008), as will be discussed in Chapter IV. Hydroxamate groups from siderophores bound to cutin-like aliphatic chains were originally suggested to be present as ligands, based on the high nitrogen and hydroxamate content (measured using a spectrophotometric method), as well as multi-dimensional NMR analysis of

the purified colloid which suggested the presence of cutin-like aliphatic structures (Xu et al., 2008).

Siderophores, produced and excreted into soils by microbes and plants for the purpose of chelating metals, form some of the strongest complexes known with iron, and are proposed to be assimilated into organic matter. Chapter IV also tests the hypothesis that a common siderophore produced by soil bacteria, desferoxamine (DFO), may react with a cutin biopolymer as a model SOM source. Very little is known about this process, however it could explain how nitrogen can be incorporated into organic matter since many of the suspected sources that persist in soils (lignin, cutin, etc.) are often devoid of it. Multi-dimensional NMR and elemental analysis are employed to evaluate reaction results, which are presented along with other potential Pu metal-binding compounds, aromatic carboxylic acids, as identified in an alkali extract of this high-Pu concentrated fraction using ESI-FTICRMS. High mass accuracy measurements of molecular formulas are also compared for the original soil organic matter, the crude colloid and the pure iso-electrically isolated colloid with the highest Pu-affinity.

The overarching goal of this thesis is to establish carboxyl containing aliphatic and aromatic compounds in soil organic matter as key components of humic acids, potentially originating from lignin and to demonstrate their significance in isolates of soil that have a strong affinity for binding plutonium and iron. Alicyclic carboxylic acids, for the first time, can be traced to lignin as a precursor, and the evidence for their cyclic structures is confirmed using multi-dimensional NMR. Further, carboxylic acids, particularly those attached to aromatic and condensed aromatic structures and containing a significant number of CHON compounds, are found in high

abundance in organic matter with high Pu-affinity. Advanced analytical techniques such as ESI-FTICRMS and multidimensional NMR are employed for this detailed molecular level analysis of natural organic matter with minimal manipulations prior to analysis that may alter its natural form.

CHAPTER II

POTENTIAL ORIGIN AND FORMATION FOR MOLECULAR COMPONENTS

OF HUMIC ACIDS IN SOILS

Preface

The content of this chapter was published in Geochimica et Cosmochimica Acta in 2016. Below

is the full citation. The formatting has been altered to incorporate the supporting information into

the body of the manuscript. See Appendix B for the copyright permission.

DiDonato, N., Chen, H., Waggoner, D., Hatcher, P.G., 2016. Potential origin and formation for

molecular components of humic acids in soils. Geochimica et Cosmochimica Acta 178, 210-222.

1. INTRODUCTION

Humification of soil organic matter has traditionally been defined as the process by which plant

and microbial debris are transformed in to humic substances (Sollins et al., 2007; Guggenberger,

2005; Stevenson, 1994). Humic substances themselves are only loosely defined as the

significantly altered macromolecules remaining in soils that are not composed of easily

recognizable known biomolecules (carbohydrates, lignin, proteins, amino-sugars, etc.). These

have further been defined operationally into three categories: base-soluble/acid-insoluble humic

acid, acid/base soluble fulvic acid and insoluble humin.

Proposed pathways for the formation of humic substances include the lignin and lignin

decomposition theories, the lignin-polyphenol theory as well as the melanoidin pathway

(Stevenson, 1994). It is generally accepted that a combination of several of these pathways with

some modifications may be responsible for producing humic substances. Also, it is well known

that physical and chemical conditions such as climate (temperature and moisture), redox conditions, pH as well as the amount and chemical compositions of organic matter source inputs determine which and how far these humification processes progress. In this paper, we refer to more 'humified' samples as those subject to these processes to a greater extent due to environmental conditions or for a longer period of time. Most commonly, though not exclusively, the extent of humification can be illustrated through changes in organic matter contained within soil horizons at increasing depth of the soil column; more 'humified', maturely decomposed material tending to be found at depth and less humified, less decomposed material at the surface (Kögel-Knabner et al., 1991). It is generally agreed that more humified organic matter is present in highly aerobic soils than in soils that experience a lower degree of aerobic decomposition (Ikeya et al., 2013).

For well over 100 years attempts have been made to structurally define the nature of humic substances and numerous models have been developed to apply a global definition of the structural entities that might constitute humic substances. Few of these models incorporate the recent views from ultrahigh resolution mass spectrometry that humic substances are generally low molecular weight (~500 Da) but extremely varied in structure (Kramer et al., 2001; Sutton and Sposito, 2005). The general picture that emerges from all the previous work and this recent work is that humic substances comprise molecules containing varying proportions of aromatic structures, aliphatic structures, proteinaceous materials, sugars, and lipids. The aromatic and aliphatic structures appear to harbor functional groups that define their solubility or insolubility in alkali and dilute acids.

The recent vision of humic substances derived from ultrahigh resolution mass spectrometry seems to suggest that thousands of molecular components exist and these molecules have some common features that allow us to class them into several types of molecules defined by the van Krevelen diagram which plots molar H/C ratios vs molar O/C ratios (van Krevelen, 1950). This type of plot was first used by Kim et al. (2003) for ultrahigh resolution mass spectral data obtained by electrospray ionization coupled to Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICR-MS). Since this first paper, hundreds of studies have employed this approach to characterize humic substances (Nebbioso and Piccolo, 2013; Sleighter and Hatcher, 2007). What is particularly salient to our understanding of humic substances is that van Krevelen plots have defined clusters of molecules that differentiate various molecular entities that serve as a basis for the current focus. The first of these represents molecules with a molecular similarity to fragments of lignin (200-800 Da), an important biopolymer component of terrestrial plants. Similarly, other regions of the van Krevelen diagram represent molecules showing a similarity to proteinaceous materials, tannins, lipids, and carbohydrates, all of which have a derivation from well-known plant biopolymers. Two other regions of van Krevelen space have been noted. One of these is a region of condensed aromatic molecules commonly associated with pyrolytic destruction of organic compounds (Kramer et al., 2004; Hockaday et al., 2006; Ikeya et al., 2013). These appear to be sufficiently adorned with oxygenated functional groups (OH, COOH, CHO) as well as N or S-containing groups to become soluble and ionizable for ESI-FTICR-MS analysis. The other region spans the van Krevelen diagram between H/C ratios of 2.0 and 0.85 and appear to be mostly aliphatic in nature (Ohno et al., 2010). However, the number of double-bond equivalents (DBE) and oxygenated functional groups associated with these molecules suggests that they include a class of molecules named by Hertkorn et al. (2006)

as CRAM (carboxyl-rich alicyclic molecules). The CRAM notation was originally used for marine-derived dissolved organic matter but it is clear that, in terrestrial dissolved and sedimentary organic matter, one observes aliphatic molecules that fit the CRAM definition in that they have multiple double-bond equivalents and contain carboxyl groups (Lam et al., 2007). In humic acids from terrestrial sources, since these molecules plot in a region beyond the CRAM region as defined by Hertkorn, we refer to them here simply as carboxyl containing aliphatic molecules (CCAM). These molecules appear to be ubiquitous components of humic substances suggesting that they might be microbial biomolecules altered by environmental processes (Nebbioso et al., 2014).

It has become clear from recent work by our group (Kramer et al., 2004; Ohno et al., 2010; Ikeya et al., 2013) and others (Grinhut et al., 2011) that humic substances in soils, humic acids in particular, show ultrahigh resolution MS data that display van Krevelen clusters of points in three principal regions: the lignin-like, the carboxyl containing aliphatic molecules (CCAM), and the condensed aromatic, or black carbon-like (BC-like) region. To demonstrate, Figure 1 shows a van Krevelen diagram for a soil humic acid whose partial mass spectral data were published by Ohno et al. (2010) , one that appears to be representative for humic acids from numerous soils distributed across the USA. The respective regions for the three classes of molecules are shown.

The current study examines humic acids from additional soil samples to further investigate the ubiquitous presence of molecular formulas plotting in the three defined regions. In addition we provide an explanation for the formation of these molecules that introduces a new perspective of the humification process. Two of the soils were collected from a single site, however, they vary

in degree of humification because one is developed on a poorly drained substrate and the other is well drained such that the organic matter accumulates under continuously oxidizing conditions. The former poorly drained soil is inundated during the wet season and is also adjacent to a small stream. Accordingly, this soil is not as humified. Other factors (pH, organic matter input, occlusion and stabilization by soil minerals/clays, etc.) are also thought to influence the extent of humification, however, in the case of the these two samples it is expected that variations in water content, aeration and sunlight play the most dominant role in this process. Our work utilizes advanced analytical techniques such as ESI-FTICR-MS and solid state NMR to more completely characterize these humic acids at the molecular level. The ESI-FTICR-MS mass spectra of a variety of other soil humic acids are also evaluated for comparison. While it is rather straightforward to explain the source of the lignin-like molecules, we propose a new process for the formation of the other two classes of molecules that compose humic acids. Because these appear to be ubiquitous across soils representing a diversity of soil environments, the process we suggest may be a global one.
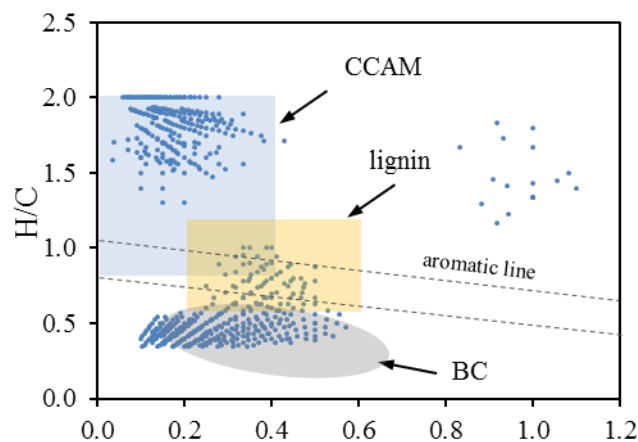
Figure 1. Van Krevelen diagram for humic acids extracted from a grassland soil in Valentine, NE. Data were obtained from Ohno et al. (2010). Only formulas containing the atoms C, H, and O are plotted. Blue, yellow, and gray regions define CCAM, lignin-like, and black carbon-like (BC) regions, respectively.

2. MATERIALS AND METHODS

2.1 Humic acid samples

Georgia humic acids were extracted from soils using 0.5 M NaOH followed by treatment with a Dowex™ ion-exchange resin to remove sodium ions. The resin was repeatedly cleaned with MilliQ water to neutrality. Humic acids were precipitated by 6 M HCl at pH 2 and freeze dried. The humic acids from Ohno et al. (2010) were treated by a process described by Olk et al. (2002) which separates two types of humic acids, those extracted without soil treatment and those extracted after acidification of the soil to release Ca-bound humic acids. We examined only the latter as they appear to be the most highly humified. The Tsubame humic acid from a buried humic layer of a highly humified forest soil in Japan was prepared as described by Ikeya et al. (2013). The Georgia soil humic acids (GA-1 and GA-2) were taken from the surface layer (top

10 cm) of a swampy hardwood forest soil and a nearby grassy, open field, respectively, in

Kingsland, GA. Armadale humic acids were obtained from the surface Ao horizon of a spodosol

soil from a deciduous forest in the Armadale region of Prince Edward Island, Canada. These

samples were provided to us by Dr. Morris Schnitzer and prior NMR studies have been

published (Kramer et al., 2001; Hatcher and Clifford, 1994; Schulten and Schnitzer, 1992; Chen

and Schnitzer, 1978). Available soil properties and characteristics are tabulated in Table 1 and

Table 2.

2.2 ESI-FTICRMS

Georgia and Armadale humic acids were dried to remove water, weighed and re-dissolved in

0.01 M NaOH overnight (~12 hours) under argon headspace. They were then batch treated with

a Dowex™ 50WX8-100 ion-exchange resin at a 3:1 v/v ratio and shaken for 1 hour to remove

Na and other cations to reduce the formation of salts upon ionization. Analysis was conducted in

Table 1

Properties of humic acid parent soils.

Sample Name	Location	Land use	Texture Class	Drainage
Valentine[a]	Cherry County, NE	Rangeland	Sand	Good
GA-1[b]	Kingsland, GA	Forest Swamp	Loam/sand	Poor
GA-2[b]	Kingsland, GA	Open Field	Sand	Good
Catlin[d]	Ogle, IL	Agriculture	Silt/clay/loam	Moderate
Armadale[c]	Prince Edward Island, Canada	Forest	Sandy loam/clay loam	Poor
Tsubame[4]	Niigata Prefecture, Japan	Sub Alpine Forest	N/A	Good

[a]USDA Official Soil Series Descriptions (OSD) website
[b]USDA Natural Resources Conservation Service Web Soil Survey
[c]Whiteside (1950)
[d]Kumada (1987),Watanabe et al. (1996)

Table 2

Properties of Soils and Humic Acids

	Soil Data				Cation Exchange Capacity (CEC)	Humic Acid Elemental Analysis		
	% Organic Matter	%C	C/N	pH	(cmol/kg or meq/100g)	% C	% N	% H
Valentine	-	0.497[a]	10.6[a]	6.0[a]	3.5[a]	-	-	-
Catlin	-	3.512[a]	13.9[a]	7.2[a]	31[a]	-	-	-
GA-1	1.0-2.0[b]	-	-	3.6-6.5[b]	2.5-7.0[b]	52.1[c]	2.9[c]	4.7[c]
GA-2	0.5-1.0[b]	-	-	4.5-6.5[b]	1.0-2.0[b]	49.9[c]	2.2[c]	3.8[c]
Armadale	3.8[d]	-	-	3.6[d]	-	56.9[d]	2.3[d]	5.2[d]
Tsubame	-	8.8[e]	23.9[e]	4.2[e]	-	55.1[e]	2.9[e]	4.3[e]

[a]Ohno et al. 2010
[b]Estimated from USDA Natural Resources Conservation Service Web Soil Survey Data
[c]This study
[d]Chen and Schnitzer 1978
[e]Watanabe et al. 1996

the negative ion mode on a Bruker Daltonics 12 Tesla Apex Qe FTICR-MS instrument equipped with an Apollo II ESI source and housed in the College of Sciences Major Instrumentation Cluster at Old Dominion University.   To improve ionization efficiency, samples were mixed 1:1 with methanol just prior to injection for a final concentration of approximately 50 ppm. A dichloromethane rinse was performed to remove organics sorbed to the Dowex resin and these samples were run separately from the aqueous samples after mixing 1:1 with methanol prior to injection. Ions were accumulated in the hexapole for 2 seconds before transfer to the ICR cell where a 4 megaword time domain was used to obtain 300 scans co-added in broadband mode using an *m/z* range of 200-1200. Spectra were externally calibrated to polyethylene glycol and internally calibrated to a fatty acid series common to natural organic matter (Sleighter et al., 2008). Empirical molecular formulas were assigned in the mass range from 200-800 m/z using an in-house Matlab code (The MathWorks, Inc., Natick, MA) according to the following criteria: $^{12}C_{2-50}, ^{1}H_{5-100}, ^{14}N_{0-6}, ^{16}O_{1-30}, ^{32}S_{0-2}$, and $^{31}P_{0-2}$ within an error of 1 ppm, and using the rules outlined by Stubbins et al. (2010). For this study, only CHO formulas were investigated, and they represent a majority (45-73%) of the total number of assigned formulas (760-1206). CHO formulas from the aqueous samples were combined with those assigned for their respective DCM extract to make a combined data set for each humic acid. Valentine and Catlin humic acids were analyzed according to Ohno et al. (2010) and Tsubame humic acids were analyzed according to Ikeya et al. (2013).

Molecular formulas were investigated using the van Krevelen diagram and categorized into three regions based on where the formulas plotted: lignin-like ($0.2 \leq O/C \leq 0.6$, $0.6 \leq H/C \leq 1.2$),

carboxyl containing aliphatic molecules (CCAM) ($0.85 \leq$ H/C ≤ 2, O/C ≤ 0.4) and condensed

aromatic (modified aromaticity index, Ai_{mod} >0.67) (Koch and Dittmar, 2006), where:

$$Ai_{mod} = \frac{1 + C - 0.5*O - S - 0.5*H}{C - 0.5*O - S - N - P} \qquad (1)$$

It is important to note that there is some overlap between the CCAM-like and lignin-like regions,

such that a small portion of molecular formulas plotting in both of these regions with

0.2<O/C<0.4, 0.8>H/C>1.2 could represent either or both of the CCAM or lignin-like categories

of molecular types based purely on their elemental ratios. For this reason, the fraction of

formulas plotting in each region does not sum to unity.

The formulas were also examined according to their Kendrick mass defect (KMD) for the

carboxyl group (COO), which is calculated as follows (Ikeya et al., 2013; Kramer et al., 2004):

$$KMD = Kendrick\ mass\ (COO) - Nominal\ Mass \qquad (2)$$

$$Kendrick\ mass\ (COO) = Exact\ Mass * (Nominal\ mass\ COO/Mass\ COO) \qquad (3)$$

2.3 NMR

Solid State ^{13}C NMR spectra were obtained on a Bruker 400 MHz Avance II spectrometer

equipped with a 4 mm solid state MAS probe. Samples were packed into an 80 μL zirconia rotor

and sealed with a Kel-F cap. Samples were spun at 11K Hz to avoid operating at the threshold of

stability for the rotor throughout the long acquisition time of the DPMAS experiment. A small

error may be associated with spinning sidebands at this frequency, however we do not expect this

to be much above the noise or exceed 1% of the signal intensity. 6400 scans were acquired

during cross-polarization magic angle spinning (CPMAS) and 2000 scans were acquired during

direct polarization magic angle spinning (DPMAS). A 1 second recycle delay and contact time

of 1 ms was used for CPMAS. Contact times were varied from 1 ms to 10 ms to find the most

appropriate contact time for magnification transfer. For DPMAS, the recycle delay was set to 50 sec to obtain relaxation of all carbons. Glycine was used as an external standard calibrated to C-2 at 42.56 ppm. Integrations were performed using the Bruker Topspin software.

2.4 Elemental analysis

The elemental composition of the Georgia humic acids was evaluated using a FlashEA 1112 elemental analyzer containing a CHN column. Samples were prepared in triplicate and calibrated to an acetanilide standard curve.

3. RESULTS

3.1 Composition of Georgia Humic Acids

3.1.1 ESI-FTICR-MS

Van Krevelen diagrams for mass spectral data from the two soil humic acids from the same region in Kingsland, GA (GA-1 and GA-2), but representing two different stages of humification are shown in Figure 2. Figure 2a shows the ESI-FTICR-MS-generated elemental data for humic acid GA-1, obtained from a poorly drained soil sampled from a deciduous forest swamp. The standing water and low oxygen content limits humification of organic matter inputs such as plant debris. In contrast, humic acids were also sampled from a nearby open field (GA-2) with grass cover. Humic acids here are exposed to more oxygen and sunlight. Regular drying of the soil allows organic matter inputs, mostly grasses, to decompose rapidly in the aerobic environment. Humic acid contents in soils can vary, and although we did not measure recoveries in this study, they can account for up to 40-50% of the organic carbon in soils.

What is immediately apparent from the data in the van Krevelen diagrams is that points appear to be concentrated in the three regions defined previously to be ubiquitously present in soils from

various other locations (Ohno et al., 2010). Points in the lignin-like region represent molecules

derived from the macromolecular lignin polymer that has undergone microbial degradation to
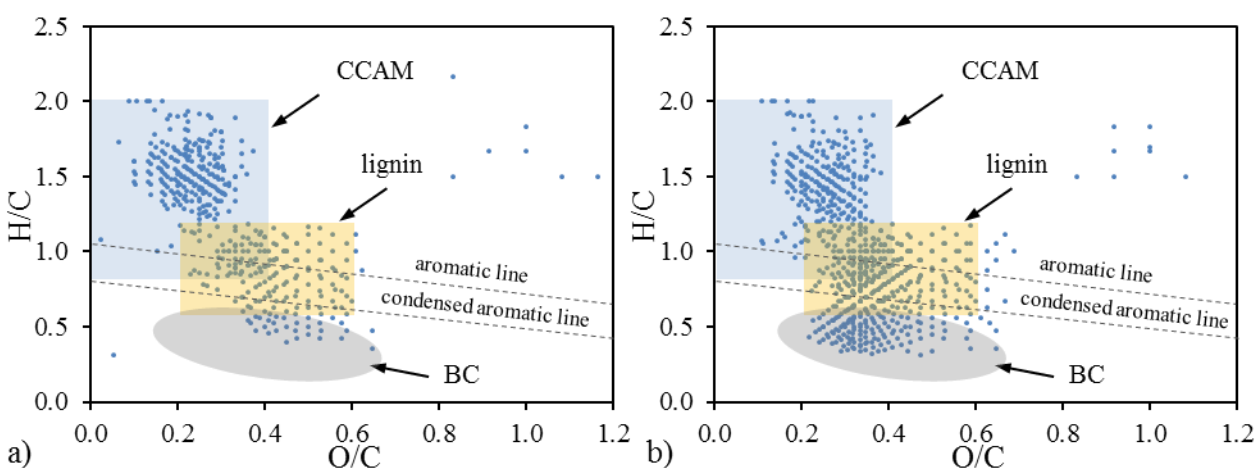
Figure 2. Van Krevelen diagrams for humic acids from a) a deciduous forest/swamp (GA-1) and

b) a grassy/open field (GA-2) in Kingsland, GA. CHO only compounds are plotted.

small molecules and has most likely been oxidized such that it becomes soluble in base to be

included as humic acids. It should be noted that other aromatic biopolymers such as condensed

tannins may also plot in this region however we do not suggest that these molecular types make a

significant contribution to these samples. They constitute a much lower percentage of total fixed

carbon in land ecosystems than lignin, which is considered second only to cellulose as the most

abundant constituent of plant biomass and makes up about 30% of the organic carbon in the

biosphere (Ruiz-Dueñas and Martínez, 2009; Boerjan et al., 2003). Hydrolysable tannins, which

plot in a region with O/C as high as 1, are either absent or poorly represented in the humic acid

ESI-FTICR-MS data, however the data does not preclude their presence within these samples. Points in the condensed aromatic region have been suggested to originate from pyrolytic processes and to be mainly substituted by carboxyl groups (Ikeya et al., 2013; Ohno et al., 2010; Kramer et al., 2004). Points in the aliphatic CRAM-like region have been suggested to derive from microbial lipids or plant terpenoids (Simpson et al., 2011; Singer et al., 2012).

Fewer molecular formulas in the lignin region of GA-1 (36% vs 42% of total peaks in GA-2, Table 3) may be due to decreased input of these molecules to soil humic acid as a result of decreased humification of the intact plant biopolymers (Kurbatov, 1968). Microbiological activity needed for chemical breakdown of these organic materials may be inhibited by the hydro-topography of this area. However, once exposed by enzymatic attack and increased humification, more molecular formulas, and a greater variety that are similar to lignin-like degradation products can be detected in humic acids as we observe in Figure 2b. With increased humification and the ensuing increased enzymatic attack by microorganisms, molecular formulas become more pronounced in the condensed aromatic region as well. This is observable in Figure 2 and Table 3 by comparing the relative number of peaks plotting in the condensed aromatic region (BC-like region; 25% vs 9% of total peaks for GA-2 and GA-1, respectively). A similar trend has been noted for humic acids from Japanese soils (Ikeya et al., 2013).

ESI-FTICR-MS mass spectra have been shown to be reproducible (Sleighter et al., 2012) and peak magnitudes may be considered semi-quantitative where response factors can be established for a majority of the molecules contained within a sample (Kamga et al., 2014). The technique has been used extensively for characterizing samples and comparing compositions based on

Table 3

Assigned CHO formulas and summed peak magnitudes for Van Krevelen Regions

Sample Name	Total No. Formulas Assigned	No. CHO Formulas Assigned	% CHO Formulas (% CHO Peak Magnitude)			% Formulas in COO series of 2 or more
			CCAM	Lignin-like	Condensed Aromatic	
Valentine	881	573	41 (78.3)	14 (3.2)	47 (16.9)	75
GA-1	925	419	67 (64.8)	36 (32.5)	9 (10.3)	55
GA-2	1206	633	50 (42.6)	42 (39.7)	25 (32)	70
Catlin	1064	775	71 (90.2)	3 (0.3)	27 (9.5)	81
Armadale	760	384	74 (76.4)	12 (8.6)	16 (14.6)	53

formula presence/absence, relative peak magnitude and the use of van Krevelen diagrams (Antony et al., 2014; Hur et al., 2010; Mesfioui et al., 2012; Reemtsma, 2009; Sleighter et al., 2009; Stubbins et al., 2010; Kim et al., 2003; Koch et al., 2008; Sleighter and Hatcher, 2008). Future work may seek to quantify absolute concentrations for individual components, but ESI-FTICR-MS is useful for our purposes to evaluate the relative abundance of compounds within a given sample and among several samples based on relative percentages of peak numbers and summed peak intensities for the van Krevelen regions of interest as shown in Table 3. NMR results also provide for a more quantitative evaluation of these samples.

3.1.2 NMR

Lignin-like compounds are also apparent from the solid state NMR spectra of GA-1 and GA-2 humic acids shown in Figure 3. Peaks characteristic of both syringyl (153 ppm) and guaiacyl (147 ppm) lignin are present, which is expected for soils from hardwood forests containing both G and S lignin such as GA-1, as well as grasslands (GA-2) which contain all three H, G and S lignin units (Koenig et al., 2010). While the peak at 147 ppm is conspicuous in both spectra, its intensity is elevated with respect to the syringyl peak in the grassland soil humic acid (GA-2). This could represent the relative contribution of lignin monomers to the soil based on the composition in the individual species present at the site, however this peak is also assigned to non-etherified guaiacyl and syringyl units of lignin and may also indicate relatively more depolymerized lignin in the GA-2 sample (Koenig et al., 2010; Chefetz et al., 2000).

Lignin ring methoxy (56 ppm) and aryl-O carbons (150 ppm) are also prevalent in these spectra

(Hatcher, 1988). The ratio of the area of these peaks (0.58) in the DPMAS spectra for the

forested area matches well with the expected lignin monomer compositions for this ecosystem,

assuming angiosperms are composed of an equal mixture of guaiacyl and syringyl monomers to

obtain an average aryl-methoxy to aryl-O ratio of 0.58.
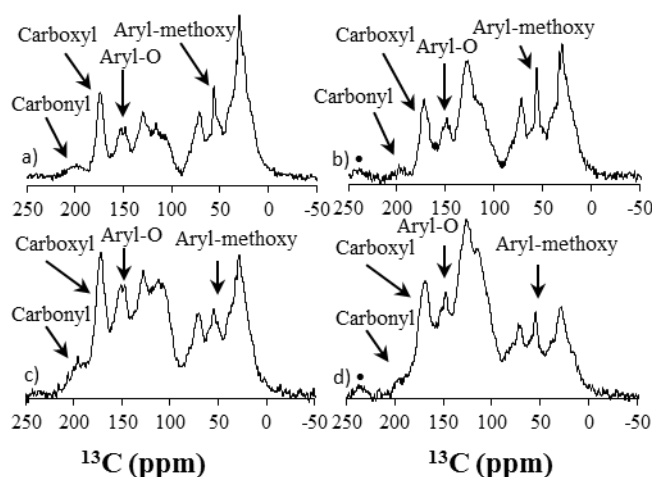
Figure 3. CPMAS and DPMAS spectra of humic acids GA-1 a) CPMAS and c) DPMAS; GA-2

humic acids b) CPMAS and d) DPMAS. Spinning side bands are denoted with a black dot.

The DPMAS spectra also reveal the high degree of aromaticity of these samples. Comparing the

aryl-O carbon peak area to the aromatic carbon peak area for these samples we notice a decrease

in this ratio for GA-2 in comparison to GA-1 (Table 4). This is in line with relatively less O-

substituted rings and potentially more humified material in GA-2.

The relative amounts of non-protonated aromatic carbons in each of the samples is also striking when comparing the DPMAS spectra with the corresponding CPMAS spectra for each sample. Because direct polarization of the carbon nuclei magnetizes all carbons equally, the spectra obtained are quantitative. In contrast, CPMAS preferentially transfers magnetization to protonated carbons, and so we can take the difference between these two spectra to estimate the amount of non-protonated aromatic, or condensed aromatic carbon in the samples. Sp^3 hybridized carbons are least affected by the difference between these two pulse programs, and so we can compare the two spectra by normalizing the integrals of each spectral region to the aliphatic region (0-45 ppm). This reveals a 16% enhancement in the area of the aromatic region (120-140 ppm) for GA-2 with direct polarization through the carbon nuclei (Table 5). In contrast, the aromatic region in the less humified sample (GA-1) that has been preserved in a poorly drained forest area contained about 12% non-protonated carbon when normalized to the aliphatic region of the CP spectra (Table 5). This is also in agreement with the ESI-FTICR-MS data, which showed a lower fraction of condensed aromatic formula for this sample than GA-2 as discussed previously in section 3.1.1. Although grasslands are prone to fire events which can contribute condensed aromatics to the soils, GA-2 samples were taken in a swampy region of southern Georgia where fires are less common. A search of the USGS Federal Fire Occurrence database, which includes data from the 1980s, showed no records of fire events in this region.

Signals from carbohydrates or alkyl-O carbons (72 ppm) are relatively weak in comparison and may also indicate the degree of humification of these samples. There is relatively less signal from carbohydrate carbons in GA-2. Likewise, the increased intensity of peaks at 168-172 (carboxyl) relative to the aliphatic region and reduced intensity of the peak at 196 (carbonyl) is in

Table 4

DPMAS Integral Areas (% distribution of carbon signal) and ratio of Aryl-O to Aromatic Carbon Signal

| | Chemical Shift (ppm) | | | | | | | |
Sample	Aliphatic (0-45)	Methoxy (45-62)	Alkyl-O (62-120)	Aromatic (120-140)	Aryl-O (140-160)	Carboxyl (160-185)	Carbonyl (185-210)	Ratio Aryl-O: Aromatic
GA-1	21.1	7.8	27.0	13.4	12.3	14.2	4.3	0.91
GA-2	15.0	6.9	31.2	20.0	11.8	13.1	2.0	0.59

Table 5

CPMAS Integral Areas (% distribution of carbon signal) and Fraction of Non-Protonated Carbon in GA-1 and GA-2 Humic Acids

| | Chemical Shift (ppm) | | Aliphatic Region Normalization Factor[a] | Normalized Area of CPMAS Aromatic Region | % Increase in Area of Aromatic Region from CPMAS |
Sample	Aliphatic (0-45)	Aromatic (120-140)			
GA-1	35.2	10.8	1.67	22.4	11.7
GA-2	25.3	17.8	1.69	33.7	15.9

[a]Ratio of aliphatic region peak areas integrated in CPMAS to that integrated in DPMAS (Table 4).

line with the sample from GA-2 containing more oxidized lignin and a higher humification potential. This is also supported by the higher percentage of molecular formulas observed by ESI-FTICR-MS in GA-2 (69% vs. 55%) to be contained within a COO Kendrick mass defect series as explained in the next section.

3.2 Comparison of Humic Acids

3.2.1 Formula Types

Soil humic acids from several other soils were also investigated in this study. The van Krevelen diagram for mass spectral data of humic acids from the Catlin soil of Ohno et al. (2010) and the Japanese soil of Ikeya et al. (2013) are depicted in Figure 4a and c. The formulas fall clearly into two predominant regions consisting of condensed aromatic molecules as well as high H/C, low O/C aliphatic CCAM molecules. Few points are observed in the lignin-like region. In contrast, the spectral data for humic acids from a spodosol soil (Armadale soil) in Figure. 4b show a less dramatic separation of these regions, with relatively more molecular formula plotting in the lignin-like region and relatively fewer condensed aromatic molecules.
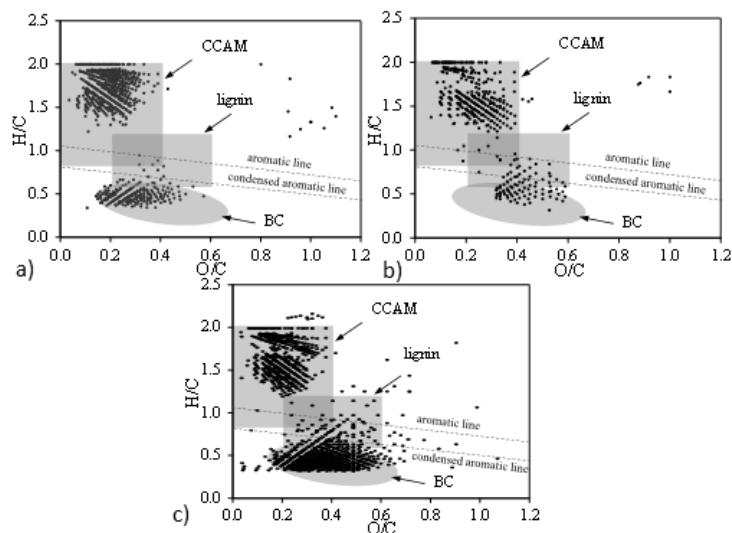
Figure 4. Van Krevelen diagrams for humic acids from various soil ecosystems and soil types. a) agricultural/native grassland in Catlin, IL, b) deciduous forest spodosol in the Armadale region of Prince Edward Island, Canada, c) the buried humic layer of a highly humified forest soil in Japan (reproduced from Ikeya et al. (2013) with permission). Only CHO compounds are plotted.

From the mass spectral observations made for the humic acids, we can readily discern a relationship based on degree of humification. The less humified samples contain relatively more components having a relationship to lignin (Kögel-Knabner et al., 1991; Ikeya et al., 2015). The more humified samples appear to mainly comprise molecules that are either condensed aromatic or are CCAM.

The range of this phenomenon is also apparent when we graph the distribution of molecular formulas in eight additional humic acid samples from the Ohno et al. (2010) study (Figure 5). These samples represent humic acids from an assortment of soil types and climate regions and should represent the diversity of soil environments across the United States. They are also considered to contain more humified organic material as a result of the isolation procedure in

which more mobile humic acids were removed. These humic acids were distinguished from their corresponding soil and plant biomass aqueous extracts by a decrease in lignin-like components and increase in both "lipid-like" and condensed aromatic components (Ohno et al., 2010). The trend is towards a collection of formulas and ion intensities primarily in the CCAM and condensed aromatic regions, with the remaining formulas plotting in the lignin-like region of the van Krevelen diagram, as can be seen in Figure 5. The average percentage of formulas and ion intensities, respectively, in the three regions were found to be: 43% ± 12% (76% intensity ± 15%) for CCAM, 48%± 13% (22% intensity ± 14%) for condensed aromatic and 9% ± 6% (2%± 2% intensity) for lignin-like regions.
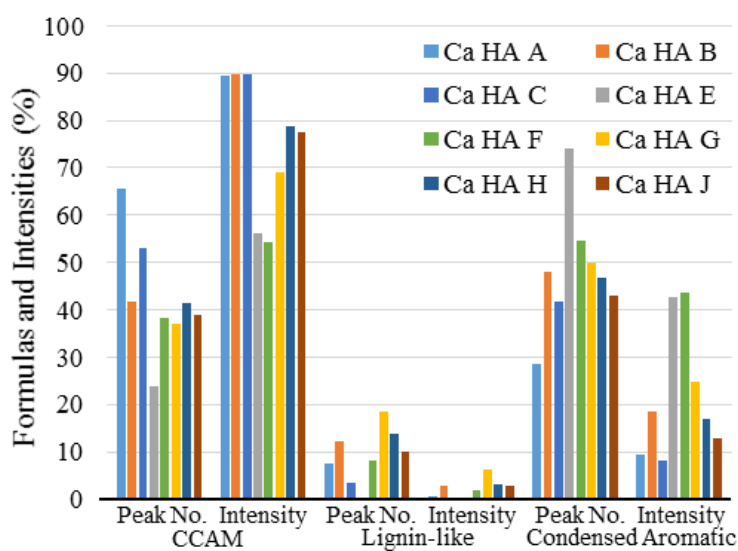
Figure 5. Relative distribution of molecular formulas and formula intensities of calcium humic acids in three identified regions of the van Krevelen diagram (carboxyl containing aliphatic molecules (CCAM), lignin-like, and condensed aromatic). Humic acids are from a variety of soil types and ecosystems at USDA sites across the US (data from Ohno et al. (2010)).

3.2.2 Carboxyl groups

While it is generally well recognized that carboxyl groups are the fundamental constituent functionalities that define humic acids, and the NMR data bear this out, we can query the elemental formula makeup of the mass spectral dataset for an indication that carboxyl groups may be components of the molecular formulas. This is accomplished by a Kendrick mass defect plot that manipulates the formula composition to align elemental formulas that differ by an exact mass of the COO group (Ikeya et al., 2013; Kramer et al., 2004). If we consider the five humic acids representing varying levels of humification (GA-1, Armadale, GA-2, Valentine and Catlin), approximately 53-81% of the molecular formulas assigned for these samples were found to be present as part of a COO KMD series of 2 or more (Figure 6), where these series extend up to six molecular formulas.
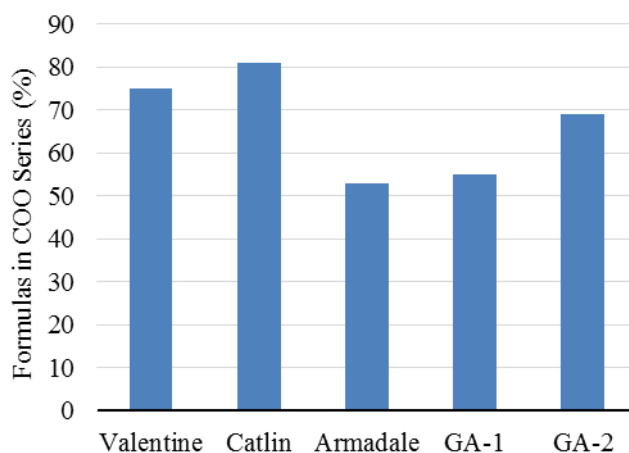
Figure 6. Percentage of formulas in humic acid samples that are part of COO Kendrick mass defect series of 2 or more.

For example, Figure 7a shows the formulas in the van Krevelen diagram for humic acids from soils at the USDA site in Catlin, IL (blue, also Figure 4a) which are contained within a COO KMD series (green, Figure 7a). These series include up to five molecular formula as shown in Figure 7b and include 81% of the CHO formulas assigned for this sample.
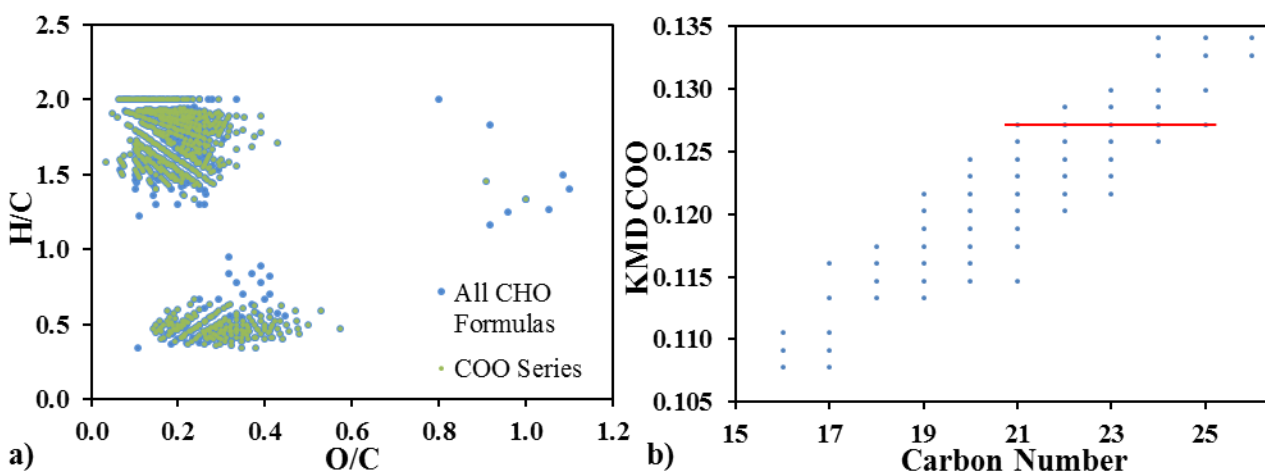
Figure 7. a) Van Krevelen diagram for calcium humic acids from a USDA site in Catlin, IL. b) COO Kendrick mass defect per carbon number for Catlin, IL humic acid molecular formulas.

When comparing the percentage of all formulas contained within a COO series among samples, those samples considered to have a higher degree of humification (Valentine, Catlin, GA-2) in general contained more formulas within a COO KMD series (75%, 81% and 70% respectively) in comparison to less humified samples (53 and 55% for the Armadale and GA-1, respectively) (Figure 6). This extends the relationship established above. The more humified samples are composed of relatively more molecular formulas in the CCAM and condensed aromatic regions

and also contain relatively more carboxylated molecular formulas than the less humified samples.

In addition, by comparing COO containing formulas for individual regions (lignin-like, condensed aromatic and CCAM) of these samples we find that molecular formulas in the regions representing more humified material (condensed aromatic and CCAM molecular formulas) in general contain relatively more COO containing formulas than the lignin-like region for each sample (Figure 8). The Valentine, NE and GA-2 humic acids are the exception, with 7 and 9% less COO KMD series formulas in the respective CCAM regions relative to the lignin regions. However, a greater percentage of COO KMD series formulas are present in the condensed aromatic regions for both samples as compared to the lignin regions (95% and 85%, respectively). Of the humic acids studied here, these two also have the lowest percentages of all formulas that fall in the CCAM region (41% and 50% respectively).

Overall, molecular formulas from the CCAM regions for these samples contained 54-82% of formulas that are part of a COO Kendrick mass defect series (Figure 8). Formulas included in a COO KMD series also accounted for 40-95% of the formulas in the condensed aromatic regions
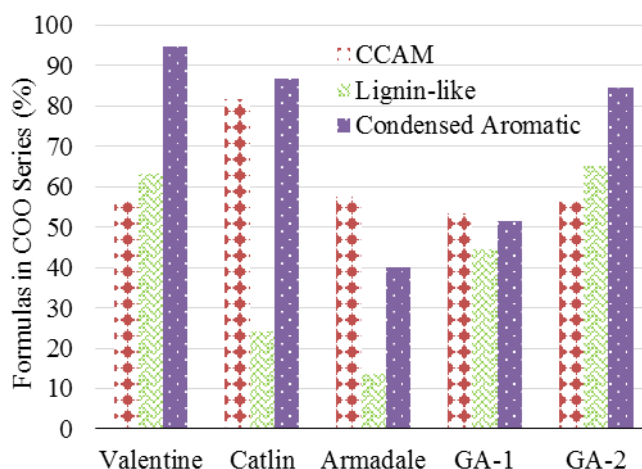
Figure 8. Percentage of formula that are part of a COO Kendrick mass defect series in each van Krevelen region of humic acid samples.

of these samples. In comparison, the lignin-like region contained relatively less, 14-65%, of formulas that were included in COO KMD series. Considering that carboxylic acids are known to be produced during lignin oxidation (Chen et al., 1983; Leonowicz et al., 2001), such a high proportion of COO KMD series formulas within the CCAM and condensed aromatic regions in comparison to the lignin region is further evidence that these regions may represent secondary, or more 'humified' molecules.

4. DISCUSSION

Soil humic acids from varying locations, soil types and degrees of humification are characterized by a common set of molecular formulas and structural entities whose presence has been debated to either derive from lignin and modified lignin, sugars and proteins, fungal or microbial phenols, or a combination of all these (Stevenson, 1994). Some recent studies in our group (Chen et al., 2014) indicate that radical polymerization reactions could play an important role in altering natural organic matter that is dominated by terrestrial lignin-derived organic matter. Chen et al.

have shown that photo-irradiation of dissolved organic matter (DOM) from the Great Dismal Swamp, Virginia, produces new particulate, base-soluble molecules of two distinct structural classes: high H/C, low O/C aliphatic (CCAM formulas) and low H/C aromatic and condensed aromatic molecular formulas with hydroxy and carboxyl substituents. The aromatic formulas, similar to black carbon from combustion sources, represent a previously unrecognized source of condensed aromatics and match formulas found in peat humic acids from the Dismal Swamp. The DOM from the Dismal Swamp is dominated by molecules derived predominantly from lignin (Hartman et al., 2015). Although primarily light-induced in the Chen et al. study, hydroxyl radicals responsible for catalyzing this oxidation are also present in soil systems. Here, fungal enzymes are one source of hydroxyl radicals (Carlile et al., 2001), others include humic substances themselves (Page et al., 2013; Page et al., 2012), and we hypothesize that reactions similar to those discovered by Chen et al. likely take place.

The details of lignin biodegradation pathways in soils can vary by organism and wood type and are not well understood, except for a few microbes and enzymes. White rot fungi are known to utilize lignin and manganese peroxidases to degrade lignin to CO_2 via several reactions, all relying on hydrogen peroxide to incorporate oxygen. Further, ring hydroxylation and opening by enzymatic hydroxyl radicals is known to produce unsaturated aliphatic carboxylic acids and hydroxylated muconic acids (Higuchi, 2004). Another recent study by our group has demonstrated similar results using Fenton reaction chemistry to produce hydroxyl radicals for the degradation of a lignin-rich wood extract (Waggoner et al., 2015). As with the study by Chen et al. the ESI-FTICR-MS results showed not only the generation of new aliphatic molecules

containing carboxylic acids, which fall in the CCAM region of the van Krevelen, but new condensed aromatic molecules as well.

Accordingly, we propose that photo-, abiotically- or microbially-generated hydroxyl radicals in soils are responsible for transforming the materials supplied to soil as fresh organic matter, mainly lignin, to the molecules observed in ESI-FTICR-MS data. When plotted on the van Krevelen diagram, the H/C and O/C ratios of molecular formulas from humic acids predictably plot in the same regions as the newly produced formulas discovered by Chen, et al. In fact, a portion of these molecular formulas were found to have the same exact composition as those identified by Chen et al. Figure 9 illustrates the photo-produced molecular formulas (red) in common with the humic acid formulas from the Valentine, NE site (blue) from Figure 1. These formulas represent approximately 16% of the humic acid CHO molecular formulas assigned for this sample.

Based on these observations, we have defined the three regions of interest for the humification process in soils as follows. (1) The region for lignin-like compounds ($0.2 \leq O/C \leq 0.6$, $0.6 \leq H/C \leq 1.2$) we propose to be the primary substrate for subsequent reactions with hydroxyl radicals in which ring opening, condensation and transformation into two distinct types of molecules occurs. These are (2) condensed aromatics ($Ai_{mod} > 0.67$) and (3) carboxyl containing aliphatic molecules (CCAM) ($0.85 \leq H/C \leq 2$, $O/C \leq 0.4$) reminiscent of carboxyl-rich alicyclic molecules (CRAM) described by Hertkorn et al. (2006).
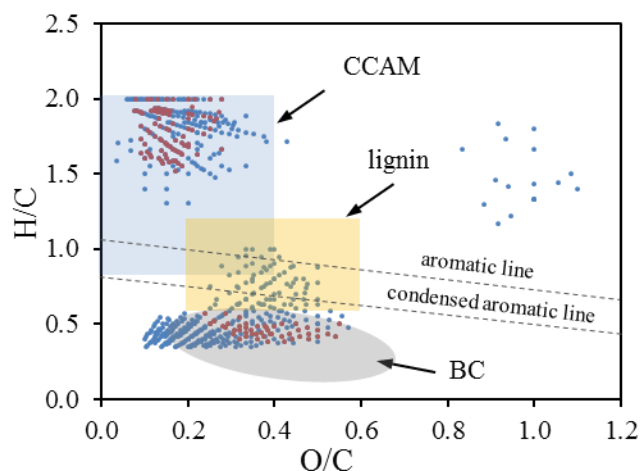
Figure 9. Van Krevelen diagram for humic acids extracted from a grassland soil in Valentine, NE (blue) with molecular formula in common with photochemically-generated formulas shown by Chen et al. (2014) overlaid in red. Only the CHO compounds are plotted.

Based on KMD COO series, carboxylic acids may functionalize the molecules that plot in these regions and we suggest this could be the result of two processes. First, the initial oxidation of lignin is known to produce carboxylic acids (Higuchi, 2004) (see also Figure 2 of Waggoner et al. 2015). These molecules would plot with higher O/C in the lignin-like or CCAM-like region of the VK diagram, as was shown in the Waggoner et al. study. Secondly, decarboxylation may occur though the loss of CO_2 during radical propagation of muconic acid-like molecules (Figure 3 Scheme, Waggoner et al. 2015) and these molecules would plot in the lower O/C and higher H/C CCAM-like region. Decarboxylation may also occur during intermolecular cyclization and radical condensation reactions (Figures 4 and 7, Waggoner et al. 2015), with these molecules plotting in the condensed aromatic or possibly the CCAM-like regions. Depending on the extent of the reactions and intermediates formed, the presence of COO KMD series within CCAM like molecules could be evidence for the first process of carboxylic acid formation as well as the

second process of decarboxylation, or both. The presence of molecules differing only by the number of carboxyl groups implies these molecules have the same core molecular formula and could be reactants or products of these processes. However, the direction of the process (carboxylation or decarboxylation) cannot be discerned from this information alone.

The modified aromaticity index (Ai_{mod}) is used here as a conservative approach to delineate the condensed aromatic region since it assumes half of the double bonds formed between C and O do not contribute to aromaticity, ring formation or condensation (Koch and Dittmar, 2006). While the actual aromaticity in a molecule can be higher, an $Ai_{mod} \geq 0.67$ provides a minimum threshold for the presence of condensed aromatic structures.

The average DBE for molecular formulas of the humic acids in this study categorized as CCAM ranges from 3.6 to 9.4, and the average Ai_{mod} ranges from 0.10-0.27, within the range for unsaturated and/or alicyclic molecules. There is evidence for the biological origin of some of these molecules based on the even over odd carbon number predominance of molecular formulas with low DBE, which suggests contributions from plant waxes (Gupta, 2014). The intensity-weighted bubble plot of DBE vs. carbon number shown in Figure 10 illustrates this most clearly for molecular formulas from C_{14} to C_{36} with DBE 1-3. These formulas most likely represent compound classes with terminal functional groups such as fatty acids, primary alcohols and aldehydes or esters (Riederer and Muller, 2008). However, the even numbered carbon prevalence diminishes with increasing DBE for all humic acid samples until there is very little preference or none at all.
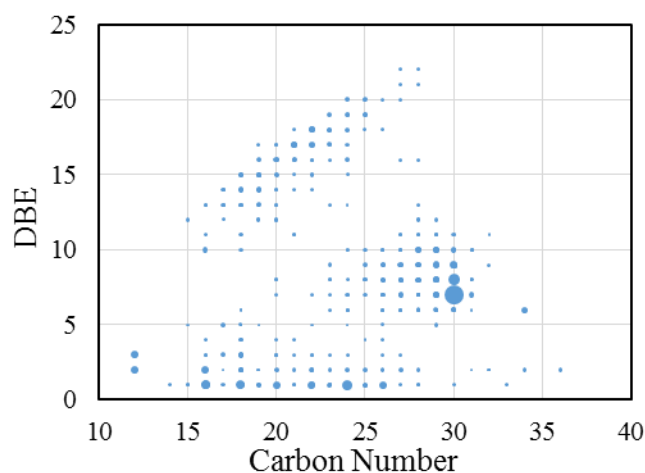
Figure 10. Intensity-weighted DBE vs. carbon number plot for the Armadale humic acid. Bubble size represents the relative individual peak intensities.

Formulas with H/C ratios approaching 2 and minimal DBEs have been shown to be created after oxidation of lignin extracts with hydroxyl radical by our group. They are also found here to be part of COO Kendrick mass defect series. Therefore, this region has been expanded from the traditional CRAM region to include formulas with H/C higher than 1.5 and slightly lower than 0.25 O/C proposed for CRAM (Hertkorn et al., 2006). Together with the condensed aromatic and lignin-like molecules, these are of interest when considering humification.

5. CONCLUSIONS

Recent application of ultrahigh resolution ESI-FTICR-MS to the study of humic substances has revealed a commonality among humic acids from varied sources and environmental conditions. In general, we can group the components of humic acids into three predominant types that, based on recent findings of organic matter-hydroxyl radical chemistry, may be inter-related chemically.

These molecular types refine the current understanding of humic acids, and may denote a pathway for the process of humification of organic matter.

Lignin-like molecules that plot in the moderate H/C and O/C region of the van Krevelen diagram may represent the first stage of humification by microbial and fungal enzymatic attack. Further reactions of these molecules with free radicals in soils, either concerted with efforts by microorganisms to access more readily available energy sources such as carbohydrates, or fortuitously due to the presence of hydroxyl radicals, can account for the remaining two types of molecules. Condensed aromatics, often containing oxygenated functional groups such as hydroxyl and carboxylic acid and previously assumed only to be present as a result of thermal processes are now recognized as a potential product of lignin humification. Similarly, CCAM and CRAM-like molecules, though not physically isolated have mostly been identified in samples from marine environments and hypothesized to be present as a result of degraded microbial biomass, are also a possible byproduct of condensation and decarboxylation reactions of lignin-like molecules in soils. The prevalence of these two groups of molecules relative to lignin-like molecules in humic acids, and the fact that many are included in COO Kendrick mass defect series also correlates well with the degree of humification.

The abundance of the biopolymer lignin and the presence of these molecular groups in far-reaching environments including the open ocean has potentially significant implications for the cycling of terrestrial organic matter. This work represents renewed evidence for lignin as a potential source of much of the organic matter in humic acids from a variety of environments.

CHAPTER III

ALICYCLIC CARBOXYLIC ACIDS IN SOIL HUMIC ACID AS DETECTED WITH ULTRAHIGH RESOLUTION MASS SPECTROMETRY AND MULTI-DIMENSIONAL NMR

Preface

The content of this chapter was submitted to Organic Geochemistry in December, 2016 and is currently under review. The formatting has been altered to incorporate the supporting information into the body of the manuscript.

1. INTRODUCTION

Several recent studies based on ESI-FTICRMS have provided evidence from elemental formulas for three predominant components of soil humic acids (Ohno et al., 2010; Ikeya et al., 2015; DiDonato et al., 2016). Molecules that resemble lignin-like formulas in their ratios of carbon, hydrogen and oxygen have been observed to be a key group of compounds, particularly in humic acid samples that are considered to be poorly humified. Two other major groups, condensed aromatic molecules as well as aliphatic molecules, the latter expected to be primarily of alicyclic nature, tend to be more abundant in more humified samples, and are thought to contain a notable proportion of carboxylic groups. Condensed aromatic structures are commonly thought to originate from thermogenic oxidation of organic matter (pyrogenic black carbon), while carboxylic-rich alicyclic structures in natural organic matter (NOM) have been proposed to have biological sources, such as microbial decomposition products in marine and coastal environments (Hertkorn et al., 2006; Hertkorn et al., 2013; Lechtenfeld et al., 2015), plant and microbial terpenoids in terrestrial environments (Lam et al., 2007), and sterols and hopanoids in

both environments (Hertkorn et al., 2006), (Woods et al., 2012). This has been primarily

construed from the detection of alicyclic structures containing abundant hydroxyl groups,

carboxylic acids and methyl groups, along with quaternary carbons in fractions of dissolved

organic matter (DOM, Woods et al., 2012).

Recent work using ESI-FTICRMS has shown that photoxidation of lignin-derived DOM (Chen

et al., 2014) and the hydroxyl radical oxidation of lignin (Waggoner et al., 2015) also produce

carboxyl containing aliphatic and alicyclic molecules in addition to condensed aromatic

molecules. Considering the fact that aliphatic carboxylic acids are known to be formed by ring-

opening reactions during transformation of lignin fragments by fungal enzymes and hydroxyl

radicals (Higuchi, 2004; Waggoner et al., 2015) it is logical to expect lignin to be the source of

these molecular types in humic acids (DiDonato et al., 2016). While it is not uncommon that

aromatic compounds in soils have been proposed as evidence for lignin as a precursor (Flaig,

1964), it is less apparent how aliphatic or alicyclic compounds could originate from lignin as

opposed to some other source (e.g., terpenoids, microbial products, etc.). The structures of

aliphatic molecules identified in humic acids and recently suggested to be derived from lignin

have been primarily deduced based on indirect measures such as molecular formula

stoichiometry, calculated numbers of rings and double bonds and related mathematical

manipulations of ESI-FTICRMS data but more detailed and direct structural characteristics of

these molecules have yet to be investigated. Accordingly, the focus of this paper is to employ

advanced multi-dimensional NMR techniques to establish the main structural motifs for soil

humic acids. We choose a humic acid sample that is typical of those found in soils. It is

important to note that previous studies by our group have demonstrated by multidimensional NMR that alkaline extracts of soils resemble the whole soil organic matter (Zhong et al., 2011).

Multidimensional NMR techniques provide a direct measure of molecular connectivity and can offer a wealth of complementary information to establish whether the postulated molecular motifs for aliphatic constituents of humic substances are alicyclic. These techniques commonly include homonuclear correlations among neighboring protons (correlation spectroscopy, COSY) and protons within the same spin systems unobstructed by heteroatoms (total correlation spectroscopy, TOCSY) as well as heteronuclear techniques that identify immediately bonded C-H resonances (HSQC, HMQC). Numerous multidimensional NMR studies of humic substances are available in the literature, mostly for DOM (Hertkorn et al., 2013; Lam et al., 2007; Lechtenfeld et al., 2015; Woods et al., 2011; Woods et al., 2012) humified plant residues (Kelleher et al., 2006; Zhong et al., 2011; Koenig et al., 2010), fulvic acids (Cook et al., 2003; Simpson et al., 2003; Simpson et al., 2001) and some humic acids (Deshmukh et al., 2007; Hertkorn et al., 2002; Kelleher and Simpson, 2006; Hsu and Hatcher, 2005; Mao et al., 2011). The presence of a variety of compounds including carbohydrates, proteins, lipids, long chain polyesters and olefins as well as phenolic and methoxy functional groups has been suggested as components of humic substances and originating from microbial and plant biopolymers such as waxes, cutin, suberin, lignin, tannins and terpenoids. A notable amount of signal from the available spectra of humic substances, humic acids in particular, can be assigned to these recognizable molecules anticipated to be present (carbohydrates, lipids, amino acids, etc.), and yet the variety of signals and chemical shift deviations from known compounds could indicate extensive alterations of parent molecules. Functional group and heteroatom substitutions appear

to be ubiquitous, and alicyclic molecules are also commonly proposed, as mentioned above, particularly for samples of marine and terrestrial DOM.

Unfortunately, there are relatively limited available multidimensional NMR spectra of humic acids in which the molecular connectivity of carboxylic acid functional groups in particular have been examined. Largely invisible to the techniques mentioned above, and limited by the low abundance of ^{13}C and lack of directly-bonded protons to transfer magnetization, more specific techniques are required to observe these functional groups. For example, the heteronuclear multiple bond coherence (HMBC) experiment is used to identify long-range interactions between ^{13}C and ^{1}H, generally up to 3 bonds. This experiment is primarily of interest for investigating the chemical environment for non-protonated carbons, specifically carbonyl carbons of carboxylic acids, which are key components of humic acids. By focusing our attention on the molecular environment in which carboxylic acids persist in humic acids, we can begin to piece together the structural characteristics that may lend insight to their source or potential transformation mechanisms. Due to the low sensitivity of this technique, however, only a handful of HMBC spectra for NOM exist (Zhong et al., 2011; Cook et al., 2003; Hertkorn et al., 2013; Lam et al., 2007; Simpson et al., 2001; Woods et al., 2012) and even fewer for humic acids (Deshmukh et al., 2007) are available in the literature. In studies where carboxylic functional groups have been specifically investigated, these have been suggested to be bound to alicyclic molecules in both marine and terrestrial DOM (Hertkorn et al., 2013; Lam et al., 2007; Woods et al., 2012) as well as soil and riverine humic acids (Deshmukh et al., 2007). In humic acids, cyclohexane carboxylic acids in plant waxes and modified cuticles have been offered as potential sources (Deshmukh et al., 2007).

In this work, we focus in more detail on the aliphatic structural entities that comprise humic acids, building on our previous work using one-dimensional NMR and ESI-FTICRMS (DiDonato et al., 2016) to assess a highly aliphatic Georgia soil humic acid in particular. This sample was found to have a predominant aliphatic component (67% of the CHO molecular formulas) yet considerable double bond equivalence (DBE) (8 on average) to suggest alicyclic rings as the framework that bears carboxylic acids. These were estimated to compose 54% of the aliphatic formulas by Kendrick mass defect analysis. Here we apply a suite of multidimensional NMR techniques (HMBC, HSQC, COSY, TOCSY) to directly assess the structural configuration of the aliphatic and carboxylic acid components, in particular, in order to offer partial validation of our previous findings. We present our results in context with available information for humic substances from the literature, in which structural entities of peat (Kelleher and Simpson, 2006; Hertkorn et al., 2002) and soil (Deshmukh et al., 2007; Kingery et al., 2000; Schmitt-Kopplin et al., 1998) humic acids have been described. The specific molecular environment of carboxylic acids as reported for fulvic acids (Cook et al., 2003), and DOM (Hertkorn et al., 2013; Lam et al., 2007; Woods et al., 2012) is also discussed for comparison.

2. MATERIALS AND METHODS

2.1 Sample preparation

Humic acids were obtained from the surface layers of a swampy hardwood forest in Kingsland, GA. Humic acids were extracted from soils as described by DiDonato et al. (2016) where the ultrahigh resolution mass spectrum and solid-state CPMAS ^{13}C NMR has been reported previously. Briefly, the soil was extracted first with organic solvents to remove lipids and then by 0.5M NaOH followed by treatment with Dowex ion-exchange resin, which was repeatedly

cleaned with milliQ water to neutrality. A small amount of 6M HCl was used to precipitate the

humic acids at pH 2, which were then freeze dried.

2.2 ESI-FTICRMS

Humic acids were previously analyzed by negative ESI-FTICRMS and the data was processed

and reported by DiDonato et al. (2016) for CHO only compounds. For detailed information

regarding ESI-FTICRMS analysis and data processing we refer readers to DiDonato et al.

(2016). For this study we utilize the previously published dataset but additionally include all

assigned formulas for heteroatom-containing (N,S,P) formulas assigned from both aqueous and

solvent-extracted data sets. Molecular formulas were categorized into three regions based on

where the formulas plotted on a van Krevelen diagram: lignin-like ($0.2 \leq O/C \leq 0.6$, $0.6 \leq H/C \leq$

1.2), carboxyl containing aliphatic molecules (CCAM) ($0.85 \leq H/C \leq 2$, $O/C \leq 0.4$) and

condensed aromatic (modified aromaticity index, $A_{imod} > 0.67$) (Koch and Dittmar, 2006), where:

$$A_{imod} = \frac{1 + C - 0.5*O - S - 0.5*H}{C - 0.5*O - S - N - P} \tag{1}$$

Kendrick mass defect (KMD) analysis was conducted for the following molecular groupings (X):

CH_2, H_2, COO, CH_3

Where

$$KMD(X) = \text{Kendrick mass }(X) - \text{Nominal Mass }(X) \tag{2}$$

$$\text{Kendrick mass }(X) = \text{Exact m/z value of peak} * (\text{Nominal mass}(X) / \text{Exact Mass}(X)) \tag{3}$$

2.3 Multidimensional liquid state NMR

All experiments were performed using a broadband inverse gradient probe fitted in a 400 MHz

Bruker Biospin AVANCE III spectrometer. The Georgia soil humic acid, called GA-1 humic

acid, was dissolved in 0.2 M $NaOD/D_2O$ at approximately 133 mg/mL and 500 uL in a 5mm

NMR tube. A one-dimensional proton spectrum was acquired with a shaped pulse for water suppression (zggpw5shapepr, NS = 2000, delay = 1 ms). One and two-dimensional spectra were calibrated to tetramethylsilane for ^{1}H and externally referenced to methanol (49.0 ppm) for ^{13}C.

Heteronuclear multiple quantum coherence (HMBC) data were acquired in magnitude mode via zero and double quantum coherence using the hmbcndprqf pulse program with water suppression (presaturation) at 320 scans per increment, a recycle delay of 1.5 s, and without decoupling during acquisition. A 50 ms delay time was used to allow for evolution of the 2-3 bond correlations, which is appropriate for coupling constants of approximately 10 Hz. The datasets were acquired with 2048 and 256 data points for the F2 (^{1}H) and F1 (^{13}C) dimensions, with spectral widths of 4006 and 27930 Hz, respectively. The data were zero filled in the F1 dimension by a factor of 2 and by a factor of 4 in the F2 dimension and processed with a sine-squared window function multiplier without line broadening. The processed spectrum was corrected for T_1 noise by calculating projections from a row in the negative ^{13}C chemical shift region where only signals from noise are expected and subtracting the projections from each row in the remaining spectrum.

Heteronuclear single quantum coherence (HSQC) spectra were acquired using the hsqcetgpprisi2 pulse program in phase sensitive mode using echo/anti-echo TPPI (time proportional phase incrementation) for gradient selection with 512 scans per increment, 1024 and 128 data points in the F2 and F1 dimensions, respectively and a delay of 1 s for T_1 relaxation. The FIDs were weighted with a sine-squared window function and zero-filled to make a 4K by 1K matrix and phase corrected without line-broadening.

The cosygpprqf pulse sequence was utilized to obtain a homonuclear ^{1}H-^{1}H correlation spectroscopy (COSY) spectrum in the magnitude mode using 2048 data points along the F2 dimension and 128 data points in the F1 dimension, with 256 transients per increment and a 1 s delay. The FIDs were weighted using a sine function in both dimensions and zero filled to 1024 and 2048 data points in the F1 and F2 dimension prior to FT and a line-broadening of 0.3 in the F1dimension.

Total correlation spectroscopy (TOCSY) was acquired using the mlevgpphw5 experiment which operates in the phase sensitive mode using watergate W5 with gradients for water suppression. A total of 2048 and 128 data points were acquired in the F2 and F1 dimensions, respectively with 256 scans per slice, 100 ms of mixing time and a recycle delay of 1second. The data were zero filled to form a 4K by 1K matrix, processed using a sine-squared window function and line broadening of 1.0 (F2) and 0.3 (F1).

Chemical shift predictions were estimated using Advanced Chemistry Developments (ACD/Labs) ACD/SpecManager 2D NMR Predictor Version 9.15. Two-dimensional NMR spectral simulations for H,H COSY and C,H COSY (HSQC, and HMBC) were obtained for proposed structures after they were entered into the ACD/ChemSketch structure drawing interface which uses algorithms based on the ACD/Labs database of known chemical structures with assigned ^{1}H and ^{13}C chemical shifts to predict spectra. Spectra, associated ^{1}H and ^{13}C chemical shifts, coupling constants and related errors for each structure were also estimated using this software.

3. RESULTS

3.1 HMBC

A variety of carbonyl groups can be observed from the HMBC spectra of the GA-1 humic acid (Figure 11). The spectrum itself is interpreted by examining the long-range ^{13}C correlations for each proton that are represented along the column of its chemical shift, as depicted by the vertical lines in Figure 11 for several proton resonances. Correlating primarily with aliphatic protons (1- 4 ppm), nearly a dozen different peaks ranging from chemical shifts of 172 to 185 ppm ^{13}C represent carbonyl carbons from different molecular environments (Figure 11 inset). Carboxylic acids cannot explicitly be distinguished from carbonyl carbons of esters and amides, which also resonate in this region and may be present to some degree in this sample. However carboxylic acids are known to be more abundant in humic acids, and previous NMR and ESI-FTICRMS analysis of this humic acid (DiDonato et al., 2016), support this. Amides are not expected to be responsible for much of the signal due to the low nitrogen content of this sample. The lack of signal for the O-alkyl portion of esters in the HSQC spectrum, as will be discussed in more detail in section 3.2, also provides more confidence for the assignment of carboxylic acids. The most intense correlation in this spectrum is from protons at 1.79 ppm correlated to carbonyl carbons at 180.3 ppm ^{13}C. Protons in this range (1.6-2.4 ppm) have been assigned to branched aliphatic structures or protons on carbons β to α-substituted alicyclic acids (Deshmukh et al., 2007). These protons also correlate with carbons in the 60-90 ppm ^{13}C region, traditionally
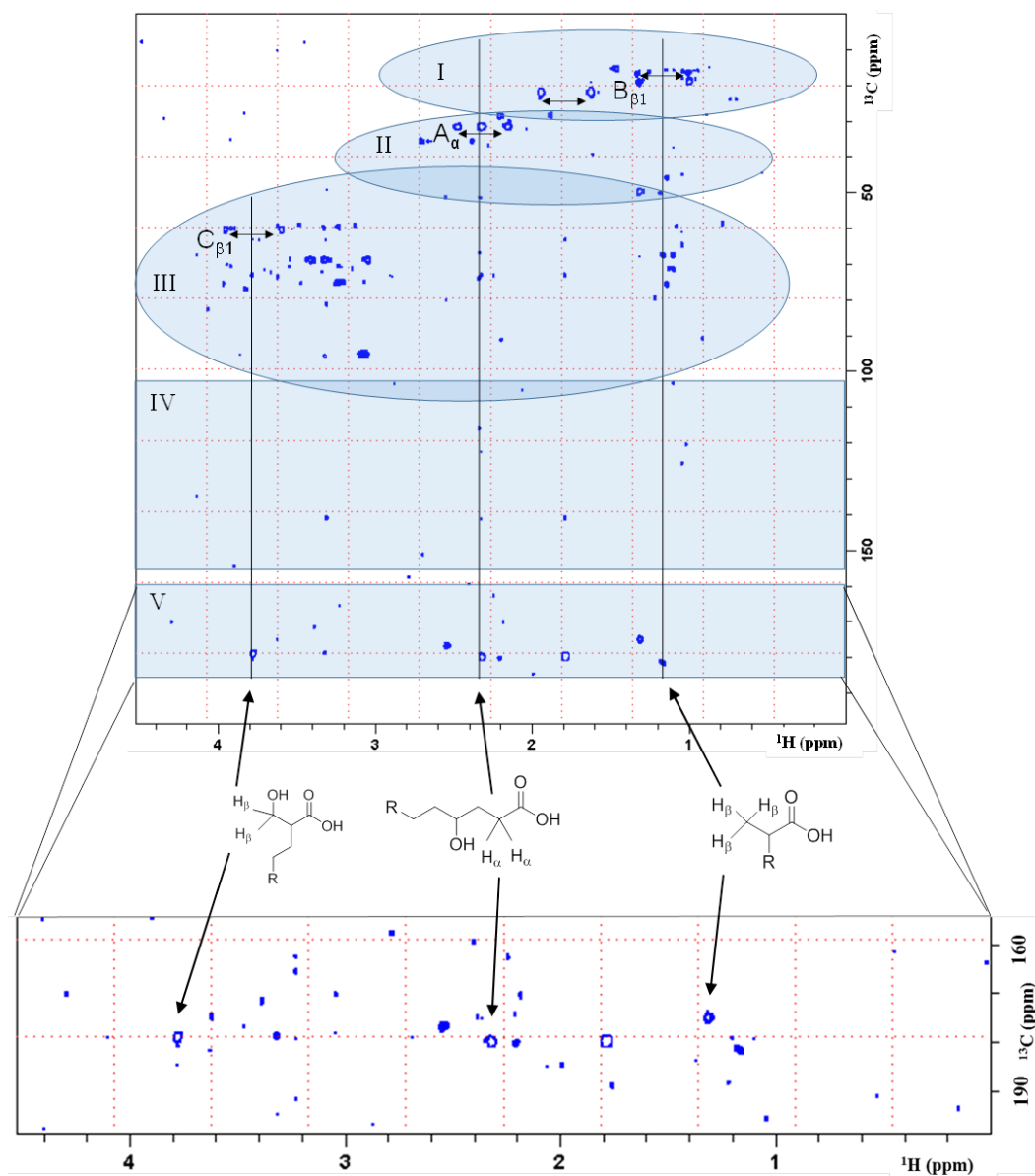
Figure 11: a) HMBC of GA-1 humic acid. Regions are defined as: I (methyl), II (methylene), III (carbohydrate/heteroatom substituted), IV(aromatic/olefin) and IV (carbonyl/carboxylic acid) b) expansion of the carbonyl carbon region. Long range ^{13}C correlations for aliphatic protons are illustrated by the vertical lines for protons at 1.2, 2.3 and 3.77 ppm which are denoted by arrows from their proposed structures (C, A and B respectively, see also Table 6). One bond

Figure 11 Continued: correlations are indicated by arrows and labeled according to the coordinates of their proposed structures in Table 6.

assigned to alcohols, carbohydrates, heteroatom substituted carbons and alicyclic molecules (region III of Figure 11). The second most intense signals are from protons at 2.3 ppm and 3.77 ppm ^{1}H correlated to carbonyls at 180.6 and 179.7 ppm ^{13}C, respectively. Protons on the carbon α to carbonyls generally resonate at 2.3 ppm, with minor variations due to substitutions or branching, and those at 3.7 ppm are often assigned to protons of heteroatom substituted carbons or functionalized carbons (OH, COOH, CHO, phenyl).

Table 6

Chemical Shift Assignments for Selected Potential Structures in GA-1 Humic Acid
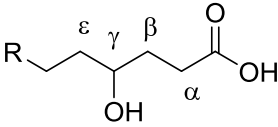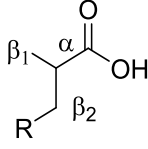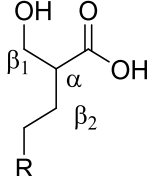
Structure	Proton Symbol	^{1}H Shift	^{13}C Shift
A	α	2.3	32
	β	1.6	30
	γ	3.5	71
	δ	1.5	40.5
B	α	2.4*	39*
	$β_1$	1.2	16.5
	$β_2$	1.7	27
C	α	2.5*	48*

Table 6 Continued
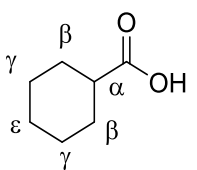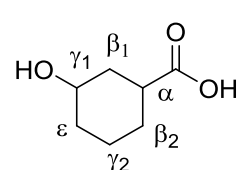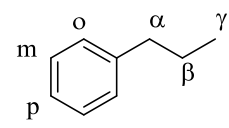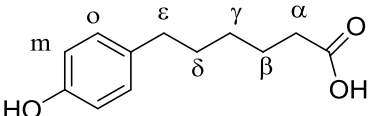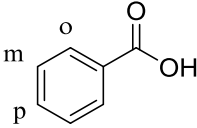
Structure	Proton Symbol	^{1}H Shift	^{13}C Shift
	β_1	3.7	61
	β_2	1.8	30
D	α	2.3 (w)	43 (w)
	β	2	29
	γ	1.6	26
	δ	1.5	26
E	α	2.7	36
	β_1	2.15	35.6
	β_2	2	28
	γ_1	3.9	68
	γ_2	1.8	20.8
	δ	1.8	35
F	α	2.6	38
	β	1.6	25
	γ	0.9	14
	o	7.2	128.7
	m	7.3	128.7
	p	7.2*	126*

Table 6 Continued

Structure	Proton Symbol	^{1}H Shift	^{13}C Shift
G	α	2.34	34
	ß	1.6	25
	γ	1.7	29
	δ	1.7	30.5
	ε	2.55	36
	o	7	129
	m	6.7	115
H	*o*	8	130
	m	7.45	129
	p	7.6*	133*

∗ shifted or not observed in spectra

(w) weak signal

Aside from the variety of protons correlated with carbonyls, one of the most salient features of this spectrum is the number of correlations of protons with carbons in the heteroatom substituted region (region III of Figure 11) commonly assigned to alcohols. In addition to the correlations mentioned above (those at 1.79, 2.3 and 3.77 ppm ^{1}H) protons at ~1.2, 1.3, 1.9, 2.2, 2.5, 3.3, 3.6 and 4.3 ppm are also observed to be correlated with carboxyl carbons and carbons in region III. For example, methylene protons (~1.3 ppm) are two to three bonds from carboxyl carbons at 175.8 ppm and oxygen substituted carbons, likely alcohols at 50.3 and 68 ppm ^{13}C. These could be attached to either straight chain or alicyclic methylene units containing oxygen substituents.

Protons at 1.3 ppm are also shown in the TOCSY spectrum to be part of the same spin system as a variety of other protons ranging from 1.4-4 ppm, as will be discussed in more detail in section 3.3.

Another important feature of the spectrum in Figure 11 is the breakthrough of several single-bond correlations. Suppression of these correlations is often achieved by including a low-pass filter or gradients within the HMBC pulse program, however many still "bleed" through in most spectra, occurring as doublets and multiplets straddling the proton chemical shift in the carbon dimension. Our attempts using pulse programs designed to suppress them were not successful, however the single-bond correlations here serve as a useful reference point for analysis, as well as a cross-reference for the HSQC spectrum in which the equivalent correlations can be observed. Notably, the one-bond correlation for α protons to carbonyls at 2.3 ppm ^{1}H is prevalent in Figure 11 at 32 ppm [13]C. This peak is also present in the HSQC spectra (discussed in the next section) and its presence here as a triplet suggests symmetrical aliphatic or alicyclic diacids. A doublet associated with the one-bond correlation for protons at 1.2 ppm ^{1}H is also visible at 16.9 ppm ^{13}C, which matches the chemical shifts for methyl substituents on alpha carbonyl carbons. Another doublet at 60.9 ppm ^{13}C spanning the proton chemical shift of 3.77 ppm is in agreement with β methylenes of carbonyl carbons on hydroxymethyl substituents of the α carbon (Table 6).

Returning to the most prevalent correlation for protons β to carboxyls (1.79 ppm), which represents a significant portion of the signal intensity, we can decipher their direct connection to carbons with chemical shifts of 22 ppm ^{13}C based on the doublet at this coordinate. Such low

carbon chemical shift is likely also due to highly shielded methyl groups, but is unusual for the relatively higher proton chemical shift than is typical for these types of methylenes.

It is interesting to note that there is little or no signal from protons β to unsubstituted straight chain and mono-alicyclic carboxylic acids (1.4-1.7 ppm). More evident are protons correlated to carboxyls between 2-2.5 ppm ^{1}H, which is in agreement with protons on substituted and cyclized carbons whose chemical shifts are more downfield than those of straight chain aliphatic carbons. This is true for the HSQC spectra as well, in which protons with chemical shifts ≥ 1.8 ppm and ≤ 1.3 ppm were most strongly observed, as will be discussed in section 3.2.

Signals at approximately ~140 ppm ^{13}C, correlating to several protons (i.e. 1.79 and 3.3 ppm) are also evident in the spectrum and could be from double bonds representing correlations between aliphatic protons and aromatic carbons, or de-shielded olefins neighboring heteroatoms or other strong electron-withdrawing groups. Peaks in this region (IV of Figure 11) are also evident in this spectrum within two to three bonds of protons at 2.3, and 3.77 ppm which also correlate to carboxyls at 180.6 and 179.7 ppm ^{13}C, respectively.

The sensitivity of aromatic carbons using multidimensional NMR techniques is notoriously low, in part due to the scarcity of protons necessary for transfer of magnetization. Very short T_2 relaxation times of longer range two and three bond couplings of rigid structures decay before they can be detected during long pulse sequences, making them more difficult to observe (Simpson et al., 2001; Cook et al., 2003). Although it is clear from the mass spectral data and one-dimensional spectra previously obtained for this sample (DiDonato et al., 2016) that these

molecular types are present, we cannot evaluate the details of their connectivity using HMBC. This technique is much less sensitive than other multidimensional NMR experiments (HSQC, HMQC, etc.). Two-bond correlations in aromatic rings in particular are very weak and often are not seen. However, the lack of signal in the aromatic region of the HMBC for carboxyls does not preclude carboxyl groups immediately bonded to aromatic carbons, due to the low sensitivity of this experiment.

While the HMBC spectrum provides some key information for understanding the molecular environment of carboxylic acids, the breakthrough of strong one bond ^{1}H-^{13}C couplings centered around the proton chemical shifts can also be a limitation requiring extra care to be taken when making assignments (Hayes and Wilson, 1997). As mentioned above, these are useful in this spectrum for several one bond correlations and do not significantly interfere with interpretation of the data, however the high intensity of these peaks may have prevented detection of more cross-peaks for long-range bonds. The peaks of this spectrum are also relatively sharp in comparison to what is expected for heterogeneous mixtures of molecules with a range of molecular size, and so resonances from smaller molecules may have been preferentially detected.

3.2 HSQC

From the HSQC spectrum we can observe a variety of structures spanning several regions of the ^{1}H-^{13}C space, including: aliphatic, functionalized and heteroatom substituted/carbohydrate, anomeric and aromatic as depicted in regions A, B and C, respectively of Figure 12. The aliphatic region (A) is dominated by C-H couplings nearby to branched or functionalized
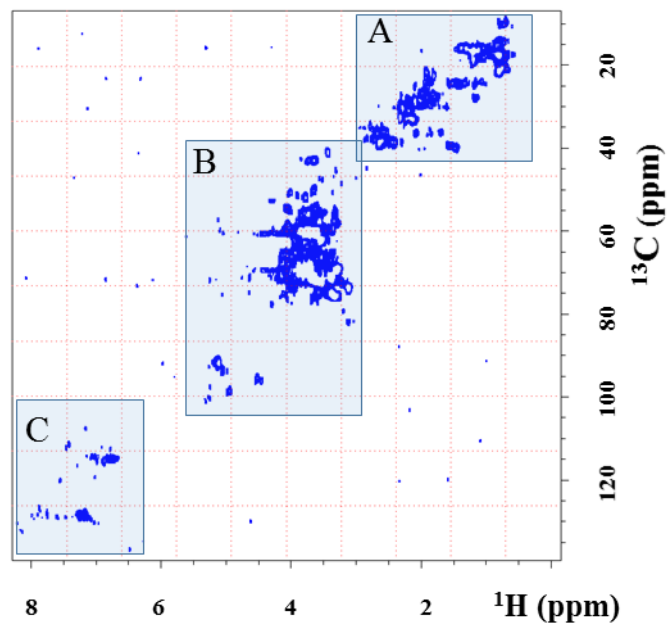
Figure 12: HSQC of GA-1 HA with designated regions A (aliphatic methyl, methylene and CH units beta to heteroatom and other functional groups), B (carbohydrate, heteroatom substituted CH groups), C (aromatic CH groups).
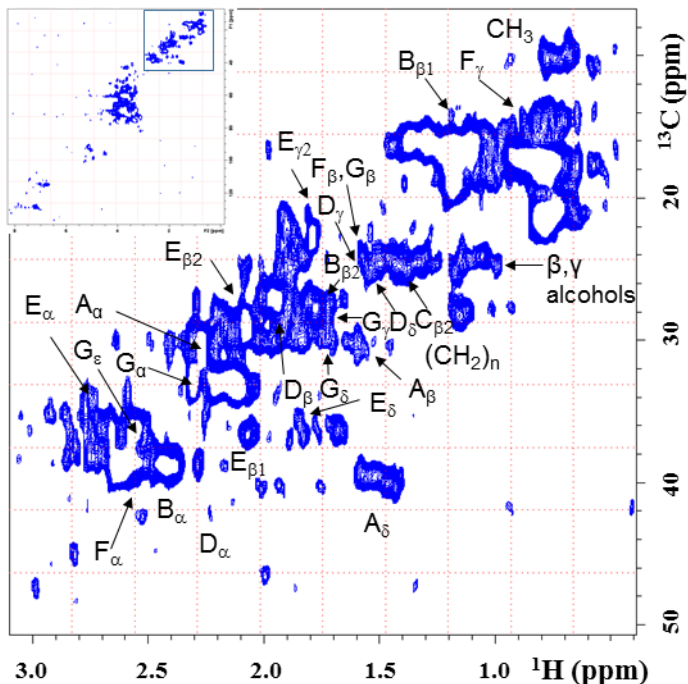
Figure 13: HSQC of GA-1 Humic Acid with aliphatic region expanded (Region A of Figure 12). Structural assignments for proposed structures are provided in Table 6.

aliphatic and alicyclic carbons (1.79-2.7 ppm ^{1}H, 22-40 ppm ^{13}C) as well as various methyl groups (0.7-1.4 ppm ^{1}H, 10-22 ppm ^{13}C). Detailed assignments are labeled in Figure 13 with associated structures and chemical shifts assignments provided in Table 6. The most intense correlation in this region resonates at 2.3 and 32.5 ppm (^{1}H, ^{13}C) which is characteristic of methylenes α to carbonyl carbons of acids and esters and is in agreement with the HMBC spectrum. Signals at 2 and 29 ppm (^{1}H, ^{13}C) are also strong and have been assigned to β methylenes of carboxylic acids on alicyclic rings. There are some signals where mid-chain methylenes could resonate (1.2, 28 ppm ^{1}H, ^{13}C) as have been observed in other fulvic and humic acids (Hertkorn et al., 2002). These are expected to originate from long chain fatty acids and esters indicative of cuticular material and plant waxes but are relatively weak in this

spectrum. They could also represent mid chain methylenes of diacids (Simpson et al., 2001). Evidence for methylenes and methines of cyclized rings is also apparent based on several peaks between 1.5-2.1 ppm ^{1}H and 35-40 ppm ^{13}C. Other relatively intense correlations occur near 2.6 and 39 ppm (^{1}H, ^{13}C), which may be attributed to alkyl substituents of aromatic rings.

As mentioned previously, it is difficult to distinguish between esters and acids using both HMBC and HSQC techniques. The α carbon of the O-alkyl portion of unsubstituted straight chain esters usually resonates around 4, 66 ppm (^{1}H, ^{13}C) and is not conspicuous in this spectrum, though could be present at 3.8, 60 ppm or 4, 54.8 ppm (^{1}H, ^{13}C) (See Figure 14). However, primary alcohols also resonate in this region and the absence of an O-alkyl peak could indicate a relatively low abundance of esters compared with acids in this sample.

Signals at 1.4, 24.5 ppm (^{1}H, ^{13}C) may represent carbons β to the carbonyl carbon; these have been reported at 1.5, 22 ppm (^{1}H, ^{13}C) for esters and closer to 24-26 ppm ^{13}C for acids (Deshmukh et al., 2003). This is not a strong signal in the HSQC, and it is not observed above the noise in the HMBC. Branching, substitutions or cyclization could shift this resonance downfield. Carbons β to the alkyl oxygen of straight chain esters are relatively weaker at approximately 1.58 and 30.5 ppm (^{1}H, ^{13}C), as are the α carbons for this side of an ester. These are also not apparent in the HMBC spectra, which despite its lower sensitivity is able to trace multiple bond correlations across heteroatoms.

Overall it appears there are few midchain methylenes as well as fewer that would be present in esters rather than acids. Methylenes attached to carbonyl carbons have been identified as well as methyl groups, most of which appear to be directly attached to O or C-H units.

The most intense peaks in the HSQC spectrum resonate in region B between 3.2-4.2 ppm ^{1}H and 40-80 ppm ^{13}C, traditionally assigned to carbohydrates (see Figure 14, and region B of Figure
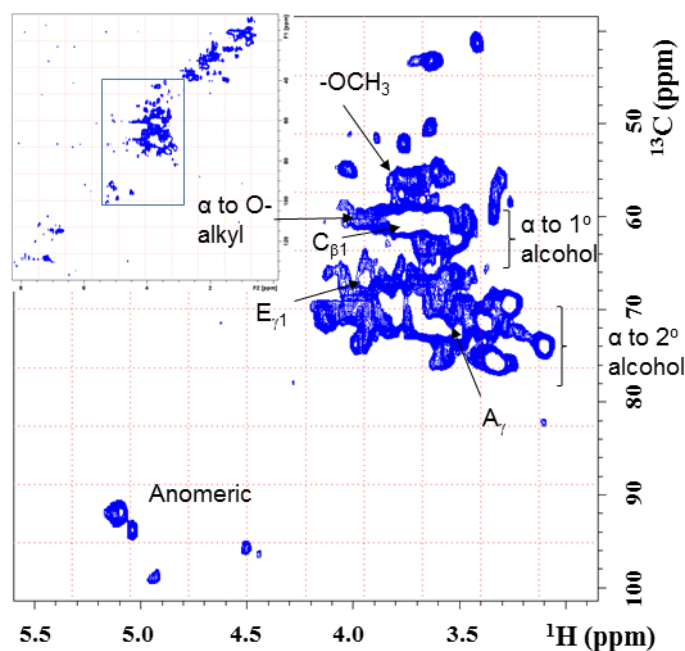
Figure 14: HSQC of GA-1 HA with heteroatom region (Region B of Figure 12) expanded. Structural assignments for proposed structures are provided in Table 6.

12). However, we do not expect carbohydrates to constitute a significant fraction of this sample, as they often range between 5-25% in soil organic matter, most of which remains in the more

soluble or fulvic fraction, with values of < 0.04% to < 6.5% reported for IHSS humic acid

standards. The solid-state ^{13}C NMR spectrum does not show a strong contribution of

carbohydrates (DiDonato et al., 2016). Here the HSQC spectrum is dominated by signals at

approximately 60 ppm ^{13}C (3.5-3.8 ppm ^1H), similar to exocyclic hydroxymethyl groups, and

alcohols and ethers at 70-76 ppm (3.1-3.8 ppm ^1H). Methoxy groups (56 ppm) are also present

but not predominant. Ring carbons of carbohydrates are often detected with these chemical

shifts, for example xylopyranoses and phenylglycosides that are known to be associated with

lignin (Yuan et al., 2011). Though not expected due to the low nitrogen content of this sample, α

and β CH groups in amino acids can also resonate in this region, for example the peaks at 3.6, 43

ppm (^1H,^{13}C) and 4.15, 71 (^1H,^{13}C) are similar to that of glycine and threonine, respectively.

However these resonances are not major contributors and those of other amino acids do not

match well with peaks in this region, nor what would be expected for their corresponding

resonances in other regions. Alternatively, we suggest that the most dominant peaks here

represent carbons on alicyclic rings adorned with hydroxyl and carboxyl functionality, consistent

with the ultrahigh resolution mass spectral data (DiDonato et al., 2016). Peaks at 63.6, 69.2, 73.7

and 68.1 ppm ^{13}C correlate with protons nearby to carboxyls (3.77 and 3.3, 2.3, and 1.2 ppm ^1H)

in the HMBC. The one bond correlation at 60.9 ppm ^{13}C and 3.77 ppm ^1H that shows up as a

doublet in the HMBC as mentioned above is also apparent here.

Anomeric carbons of oligo/polysaccharides appear at 5.1 and 92 ppm (^1H, ^{13}C), 5 and 94, 4.5 and

96 and 4.9, 99 ppm (^1H, ^{13}C); the signal at 5.1 and 92 ppm (^1H, ^{13}C) being the most intense,

albeit less intense than the heteroatom substituted region described above, yet comparable in

relative intensity to that of the functionalized aliphatic region (2-2.7 ppm ^1H, 30-40 ^{13}C). Peaks

in these regions are also in general agreement (with some deviation from published values due to modifications that may occur in humics) with that of fucose, glucose, ribopyranose, galactose assigned by Hertkorn et al. (2002) or those Yuan et al. (2011) assigned to xylopyranose, monosaccharides and phenlyglycosides and suggest at least some carbohydrates remain intact in this humic acid even though the solids NMR spectra suggest they are minor components (DiDonato et al., 2016).

Signals in the aromatic region are slightly less intense than other regions of the spectrum with two primary clusters of cross peaks (Figure 15). At approximately 128-129 ppm ^{13}C and between 7.3-8 ppm ^{1}H these could indicate the presence of COR groups in the form of ester or ethers directly attached to aromatic carbons (Perdue et al., 2007) as well as alkyl substituted rings. Ring carbons of *ortho* substituted carboxylic acids could also be included in this region, often at chemical shifts >7.5 ppm ^{13}C due to their strong proton deshielding effect (Hertkorn et al., 2002). This is likely considering several carboxyl KMD series for low H/C compounds were detected in the ultrahigh resolution mass spectral data (DiDonato et al., 2016). Resonance signals for *ortho* and *para* oxygen substituted groups are also observed at 115-116 ppm ^{13}C, and 6.8-7.3 ppm ^{1}H. These have also been assigned to guaiacyl units of lignin and both chemical shift regions match those of ring carbons for *p*-coumaryl alcohol. As mentioned above, sensitivity in this region may be hindered due to fewer protonated carbons and shorter relaxation time due to rigid structures. Aside from the mass spectral data, both HMBC, TOCSY and COSY
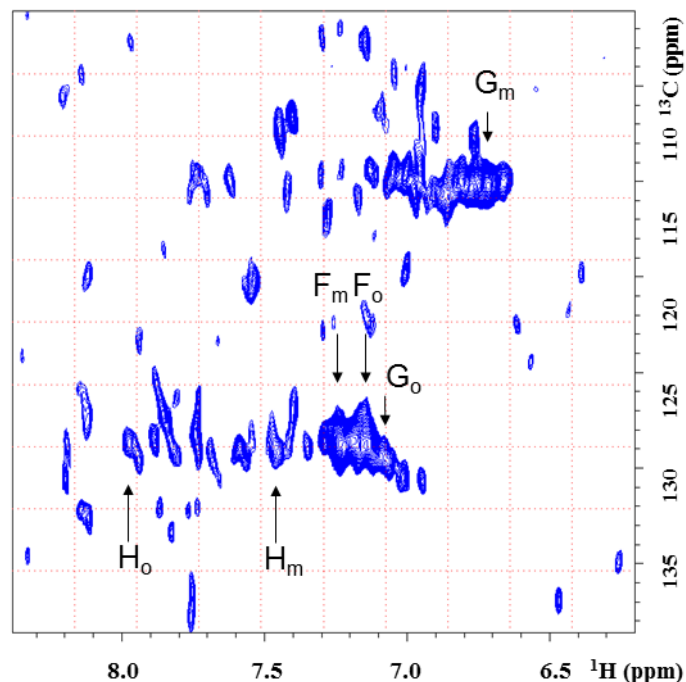
Figure 15: HSQC of GA-1 HA Aromatic region (Region C of Figure 12). Assignments are labeled according to the coordinates for their proposed structures listed in Table 6.

techniques did not provide any further diagnostic information regarding this region and so it is difficult to conclude which structures are most represented in this sample.

3.3 TOCSY

Total correlation spectroscopy (TOCSY) does provide more information concerning the aliphatic region of this sample. TOCSY is a technique that reveals which proton resonances are within the same spin systems. Terminal methyls (0.89 ppm) show coupling with branched, functionalized or cyclic methylene protons (1.14, 1.34, 1.5, 2.0 ppm) and protons of heteroatom substituted methylenes at 4.0 ppm (Figure 16). This coupling shows that an ample number of methyl groups are not attached to quaternary carbons, as has been found for some fractions of DOM (Woods et al., 2012), since spin-spin couplings detected by this technique cannot transmit across quaternary
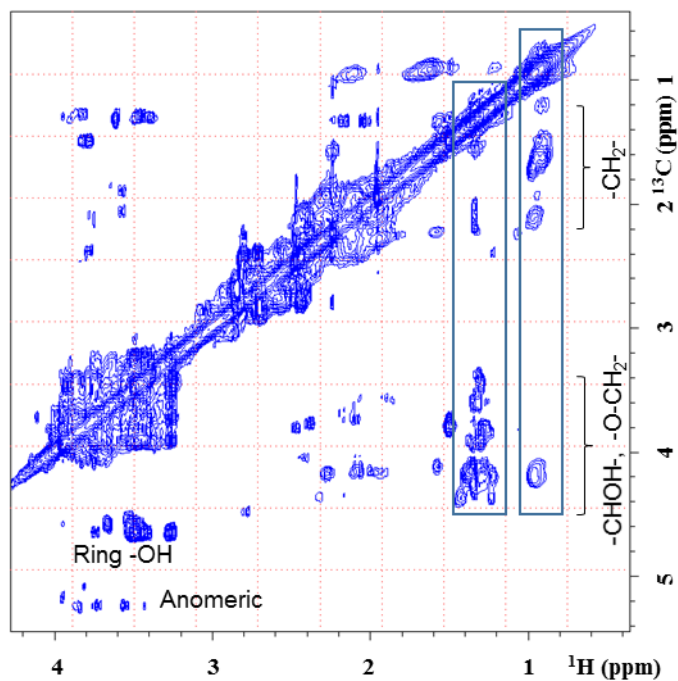
Figure 16: Total correlation spectroscopy (TOCSY) spectra of GA-1

carbons or heteroatoms. However, coupling to protons with chemical shifts of 4.0 can be accounted for by cyclic and highly branched structures with ester or alcohol functionality. Methylene protons at 1.3 ppm also couple with those at 1.4, 1.9 and 2.0 as well as protons at 3.5, 4.0 and 4.15 ppm. These protons (1.3 ppm) were also shown to correlate with carboxyl groups as well in the HMBC.

Overall the combined evidence for at least two spin systems uniting a range of protons in TOCSY, strong signals from substituted or cyclic CHs at the sacrifice of signals from mid chain methylenes in the HSQC and a variety of carbons within 2-3 bonds of most of the carboxyls detected in the HMBC, makes sense for highly branched, substituted or cyclic acid structures.

Additional information concerning the molecular makeup of individual compounds that contribute to these spectra can also be obtained from the ESI-FTICRMS data presented by DiDonato et al. (2016) and further investigated in this study.

3.4 ESI-FTICRMS

From the ESI-FTICRMS data alone, we can only indirectly predict structural assignments based on H/C ratios, calculated DBEs and aromaticity indices. However, in combination with the information from multidimensional NMR studies, we may begin to sketch out a more complete picture of the molecular environment of this humic acid.

Molecular formula identifications and carboxyl KMD analysis have been previously reported and carboxyl KMD series were found to constitute 55% of the assigned CHO-only formulas for GA-1 (DiDonato et al., 2016). Approximately 44% of the lignin-like, 55% of the condensed aromatic, or black carbon (BC)-like, and 54% of the carboxyl-containing alicyclic molecule (CCAM) formulas were part of a COO series of two or more. These regions represented 36, 9 and 67% of the CHO compounds, respectively (note these regions do not sum to unity due to overlap). For this study we incorporate the remaining formulas into our analysis, which include the following molecular formula types and percentage of formula makeup: CHO (50%), CHOS (19%), CHONS (12%), CHON (6%), CHOPN (6%), CHOPS (3%), CHOP (4%). Approximately 38, 26 and 41% of the total formulas are contained within the lignin-like, condensed aromatic and CCAM regions, respectively (Figure 17). The largest percentage of heteroatom containing formulas not including CHO compounds were observed in the condensed aromatic region, constituting 83% of the formulas in this region.
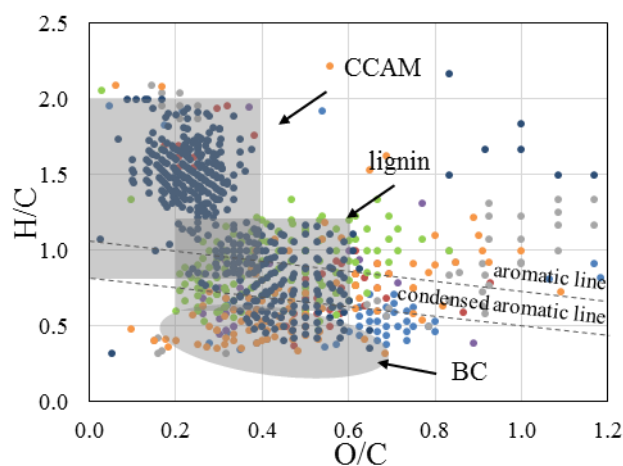
Figure 17: van Krevelen diagram of all assigned formulas in GA-1 humic acid. Dark blue = CHO (50%), green = CHOS (19%), light-blue = CHON (6%), orange = CHONS (12%), gray = CHOPN (6%), red = CHOP (4%), purple CHOPS (3%).

About 37% of total formulas were contained within a COO series in comparison to 55% of CHO only compounds contained within COO series found previously. This is largely due to dilution of the CHO formulas with the remaining heteroatom containing formulas which primarily were not part of COO KMD series. Lignin-like, condensed aromatic and CCAM molecules each were composed of 36, 18 and 45% of formulas in COO series, respectively. Since most of the HMBC carboxyl signal appeared to be tied to aliphatic carbons, we further investigated the molecular formula makeup of peaks contained within a COO series in the CCAM region to inform our structural predictions. We found an average number of 27 ($\pm$ 3) carbons, 41 ($\pm$ 6) hydrogens, 6 ($\pm$2) oxygens, and 8 ($\pm$2) DBE per molecule for CCAM molecular formulas within a COO series of 2 or more (Table 7). These were all found to be CHO compounds except one group of CHOS

compounds. Also evaluated were the number of formulas that were contained within CH_3 series, since the signal from these groups in the HSQC appeared to be relatively intense. We found 49% of the total formulas were included in CH_3 series of 2 or more, with series up to 7 molecular formulas. In addition, examining only the CH_3 series that also were part of a COO series, we found these formulas accounted for a quarter of all formulas. Of just the COO series formulas, about 67% also were part of a CH_3 series. These CH_3 series that also are part of a COO series additionally accounted for 37% of all CCAM molecules, and these formulas that fall in the CCAM region also make up 61.7% of all the formulas that are contained within both series types. Formula compositions for each of these groupings are listed in Table 7.

Table 7: Statistical CCAM formula compositions by grouping

	CCAM				CCAM COO				CCAM COO CH_3			
	C	H	O	DBE	C	H	O	DBE	C	H	O	DBE
Avg	26 ± 5	37 ± 12	6 ± 2	8 ± 3	27± 3	41± 6	6 ± 2	8 ± 2	27 ± 4	39 ± 9	7 ± 1	9 ± 2
Min	13	13	1	1	17	16	2	1	16	14	4	4
Max	39	71	9	19	32	52	9	14	32	52	9	14
Median	27	38	6	8	28	42	6	8	28	40	7	9
Mode	29	40	7	9	29	44	6	8	30	44	7	8

Based on the molecular formula information for molecules plotting in the CCAM region ($0.85 \leq$ H/C ≤ 2.0, O/C ≤ 0.4) and the structural motifs identified from the NMR data (Table 7), hypothetical structures that may be formed during the radical oxidation of lignin were developed using mechanisms adapted from Higuchi (2004) and Waggoner et al (2015). Lignin dimers and tetramers linked through common bonds, β-1 and 5-5, were used as starting material (Figure 18).
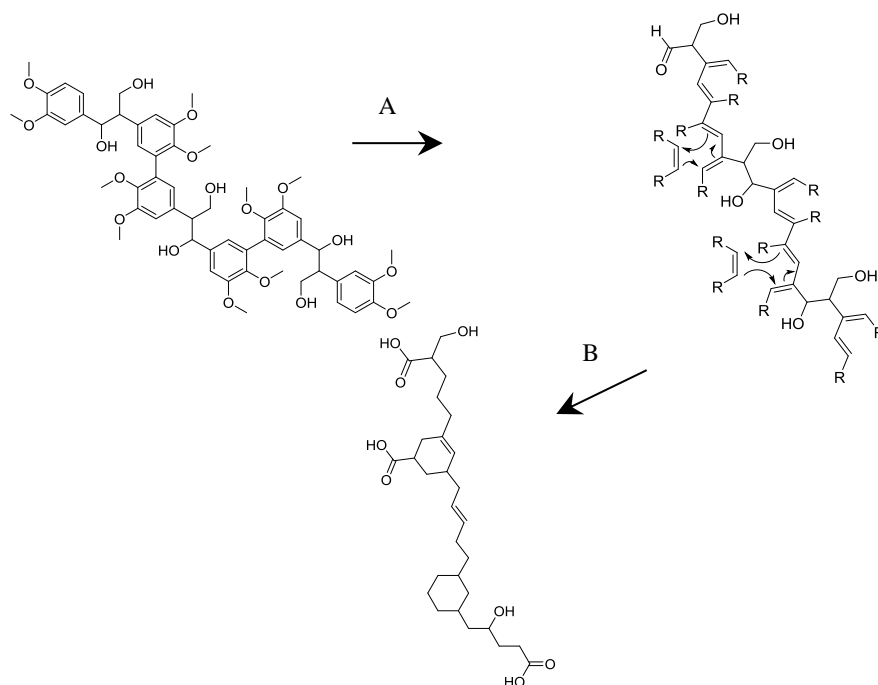
Figure 18: Hypothetical pathway for the formation of alicyclic molecules from lignin (adapted from Waggoner et al.(2015), Higuchi (2004), R=COOH). This includes 1) oxidation, depolymerization, demethylation and ring opening of a ß-1, 5-5 lignin polymer as well as alkyl phenyl cleavage of a phenolic ß-1 lignin substructure and 2) decarboxylation, oxidation of open-ring structures, electrocyclization with maleic acids, aldehyde oxidation, radical alkylation and oxidation or radical addition to double bonds.

These represent 18-25% and 7-10% of linkages in soft wood, respectively (Argyropoulos et al., 2002; Adler, 1977; Heitner et al., 2016) and were utilized in this study due to the similarity of the chemical shifts for these bonds with the observed spectra. Signals from the more common β-O-4 ether bonds are not as evident. Ring opening initiated by proton abstraction via hydroxyl radical, a process which has also been shown to form cis-cis muconic acids, followed by re-cyclization with maleic acid (a major decomposition product of muconic acid (Devlin and Harris, 1984; Shende and Levec, 2000) and strong dienophile) could lead to the central alicyclic ring structures

illustrated in Figures 19a-c. Peripheral aliphatic appendages could be remnants of the precursor lignin side-chain, opened ring, or maleic acid linkage. If radicals are formed either via reactive oxygen species or during loss of CO_2, alkyl chain addition to the ring or double bonds may also be possible.
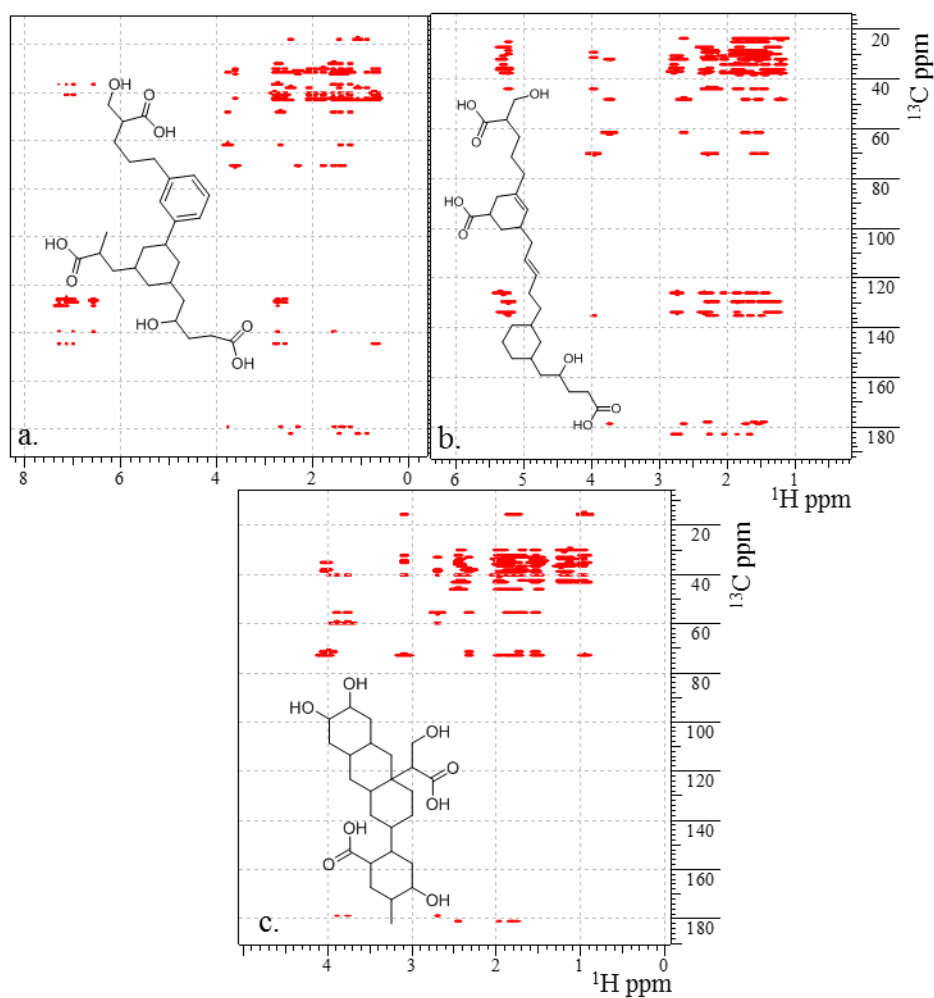
Figure 19: ACD labs HMBC Spectral Prediction for model alicyclic molecules derived from lignin a) $C_{27}H_{40}O_8$ (FW 492.6017, DBE=8) b) $C_{29}H_{44}O_8$ (FW 520.6549, DBE=7) c) $C_{25}H_{40}O_8$ (FW 468.5803, DBE=6)

Spectral predictions were conducted for alicyclic molecules ranging in rings from 1-6, from 1-3 carboxyl groups and from 1-5 alcohols. Predictions for straight chained fatty acid and ester type molecules containing hydroxyl substituents were also conducted for comparison. Both HSQC and HMBC spectra were predicted to maintain consistency. Figure 19a-c presents three of the proposed structures and their predicted HMBC spectra.

The range of correlations for carboxyl cross peaks matches fairly well with the data observed in the spectrum for GA-1, however there is noticeably less signal in the region between 55-70 ^{13}C and above 3 ppm ^{1}H in the predicted spectra of Figure 19a and Figure 19b than in the observed humic acid spectrum which is more similar to Figure 19c in this region. This may be due to additional molecules with hydroxylated rings or carbohydrates that may be present in the sample. A greater variety of signals less than 2 ppm ^{1}H and between 40 and 50 ppm ^{13}C in the predicted spectra in comparison to the observed spectrum for GA-1 may be due to weaker correlations associated with poorly resolved, complex proton multiplets in the sample (Claridge, 2009). Signals from methylenes in rings or chains that would plot in this region will be attenuated below the signal to noise threshold by complex coupling patterns. The stronger single bond correlations that have not been filtered out may also obscure the signal. Despite these inconsistencies, the model predictions in Figure 19a and b for carboxyl ^{1}H-^{13}C correlations at 1.2, 2.3, 3.7 ppm ^{1}H in particular are in strong agreement with our results. Carboxylic acid β protons with methyl (1.2 ppm ^{1}H) and hydroxyl methylene (3.7 ppm ^{1}H) substitutions on α carbons, as well as carboxylic acid α carbons (2.3 ppm ^{1}H) with alcohol substitutions on the γ carbon in which the corresponding single-bond ^{13}C chemical shifts are known, match both the HMBC and HSQC data.

Considering the stoichiometric constraints set by the mass spectral data (Table 7), which limits the number of oxygens in particular, and thus alcohols, carboxylic acids, ethers, and other oxygenated groups per molecule, it is more likely that a large variety of molecular configurations are present in the sample to account for the range of chemical shifts observed in the HMBC spectra such that these cannot be wholly represented by one or two model compounds. Nevertheless, the above structures depict some of the most significant features that can be ascertained from both the mass spectra and NMR data, namely the configurations to match the detected molecular size, stoichiometry and double-bond/ring equivalents as well as the observed chemical shifts.

4. DISCUSSION

4.1 Comparison of structural findings

Our investigation thus far of this highly aliphatic humic acid supports the occurrence of alicyclic carboxylic acid structures as components. The most notable features that can be discerned from this study are the strong resonances in the carbohydrate region, various methyl resonances and assortment of aliphatic proton correlations to carboxylic acids that are also linked to methyl, hydroxyl and olefinic carbons. Methoxy and aromatic resonances suggestive of lignin are present, but relatively minor, as are straight chain methylene carbon units expected for paraffinic material and commonly identified in humic materials (Schnitzer and Khan, 1972).

The aliphatic portion of humic acids have been well-studied and suggested to originate from biomolecules such as proteins, carbohydrates, lipids, fatty acids and waxes. Of the humic acids that have been studied using non-destructive multi-dimensional NMR techniques, paraffinic materials appear to be common especially in peat humic acids (Hertkorn et al. (2002); (Kelleher

et al., 2006), which are expected to be less humified than soils. The main difference between the

Georgia soil humic acid spectrum and other soil humic material is the less prominent signal from

methylene carbons in long chains, as well as for esters. This is perhaps due to the fact that the

soil sample was extracted with methylene chloride prior to alkali extraction. In contrast, the

Georgia soil humic acid rather contains more signal from a variety of methyl groups observed in

the HSQC spectrum between 10-25 ppm ^{13}C and 0.5-1 ppm ^{1}H as well as chemical shifts for

single and double heteroatom substituted carbons which suggest highly branched and cyclic

molecules more similar to carbohydrates or proteins. Methyl resonances are not unique to this

sample, however, and have been identified in peat and soil humic acids as well. In general, these

similarities are evident in spectra of soil humic acids published by Schmitt-Kopplin et al. (1998),

Kingery et al. (2000) and Deshmukh et al. (2007) as well as alkali soil extracts and peat HA

studied by Kelleher and Simpson (2006) and Hertkorn et al (2002). In the Hertkorn et al. (2002)

study, these were found to be more evident in the peat HA in comparison to its FA, and resemble

the same chemical shifts for methyl groups of amino acids I, V, L, A and T.

The low nitrogen content of the Georgia soil sample and ambiguous assignments for remaining

amino acid correlations suggest these compounds, if present, would not exist in this sample in

their free form. In addition, mass spectral data show that CHON compounds constitute only 6%

of the observed molecular formulas and plot predominantly in the condensed aromatic region, as

opposed to the region just above the lignin-like region where proteinaceous material is typical

(Figure 17). Other nitrogen containing compounds, CHONS constituting the largest grouping

(12% of the formulas), span a large O/C range, from condensed aromatic and lignin-like to

tannin-like, which suggests these structures, if present, could be randomly incorporated in the

macromolecular structure of organic matter. Formulas within CH_3 KMD series primarily consist of CHO (81%) or CHOS (16%) compounds and are also not limited to one region of the van Krevelen diagram, but distributed somewhat uniformly within the CCAM, lignin-like and aromatic regions. The methyl groups observed in the HSQC spectrum could therefore occur within varied chemical environments of alicyclic molecules or in side chains of degraded lignin molecules which, depending on the length of the side chain, heteroatom stoichiometry and number of attached rings could fall in either of these regions.

Corresponding resonances for carboxylic acid groups in both the HMBC and HSQC confirm correlations between carboxyls and methyl, methylene and heteroatom substituted aliphatic carbons and demonstrate the plausibility of cyclic structures. A strong correlation for alpha protons (2.3 ppm) of carboxylic acids where the intensity is less relative to more upfield chemical shifts (1.79 ppm) may also suggest branched aliphatic and alicyclic structures to be more prevalent in this sample. Deshmukh et al. (2007) used the same rationale in their study of Suwannee River and Elliott Soil humic acids, with respect to protons at 1.9 ppm as compared to the respective fulvic acid samples where the 2.3 ppm protons dominated, instead suggesting more straight chain carboxylic acids. The high DBE for aliphatic molecular formulas (average 8, mode of 9) also serves as evidence for alicyclic structures, specifically for molecular formulas that fall in both COO and CH_3 KMD series for which the minimum DBE was found to be 4 (max 14) with an average of 9 DBE (mode of 8 DBE).

We compared these findings with other HMBC spectra for humic substances in the literature, although data is limited for humic acids as mentioned in section 1, more studies are available for

fulvic acids (Simpson et al., 2001; Cook et al., 2003) and for DOM from lacustrine (Lam et al., 2007), riverine (Woods et al., 2012) and marine (Hertkorn et al., 2013) environments. Most studies of carboxylic acids in terrestrial soil (Laurentian fulvic acid (LFA), Elliott soil humic acid), riverine (Suwannee River humic acids and DOM), and even marine UDOM (Hertkorn et al., 2013) organic matter share many of the same features. One exception is for Lake Ontario DOM (LODOM) which did not indicate the same correlations of carboxyl groups with methylene chains nor were carboxyl groups in the CRAM region of LODOM found to correlate with other functionalities (methyl, hydroxyl carbons, double bond) (Lam et al., 2007). The 140-150 ppm ^{13}C region where double bonds would resonate is also not as apparent in the Suwannee River and Elliot soil spectra (Deshmukh et al., 2007), but are more evident in the LFA analyzed by Cook et al. (2003), the DOM of Hertkorn et al. (2013) and Woods et al. (2012). In contrast there is relatively less signal in the 80-90 ppm range of GA-1 where Lam et al. (2007) identified oxygenated functional groups in the MDTL (material derived from linear terpenoids) region of LODOM, which they propose to result from hydration of conjugated double bonds in these structures. The carboxylic groups in the GA-1 HA sample resonate at a slightly lower field (175.6-179 ppm ^{13}C) than the Suwannee River and Elliot HA of the Deshmukh et al. (2007) study (168-177 ppm ^{13}C) as well as the LFA of Cook et al. (2003). These chemical shifts are more similar to the Elliott FA and to the region identified as material derived from linear terpenoids (MDLT) by Lam et al. (2007) that resonate as high as 180 ppm ^{13}C.

4.2 Conclusions

Most recent studies suggest a primarily biological source for alicyclic molecules (microbial DOM, terpenoids, etc.) in humic substances, although other explanations are also possible. Hertkorn et al. (2002) attributed some of the peaks in the aliphatic region of peat HA (20-40 ppm

^{13}C, 1-2.5 ppm ^1H) to terpenoid-type structures, as well as lipids, fatty acid derivatives and aliphatic residues of deoxy sugars. Cook et al. (2003) found mostly aliphatic carboxylic acids and proposed aliphatic hydroxy carboxylic acids of alkanoic and benzoic nature, both cyclic and straight chain, as well as glucosides derived from terpenoids and flavonoids as model structures for Laurentian fulvic acid from the B_h horizon of a boreal podzol. Woods et al. (2012) also proposed cyclic structures when interpreting the HSQC and HMBC spectra of a fraction of Suwannee River DOM (SRDOM), based on the absence of mid-chain methylenes and a wide distribution of carboxyl groups within 3 bonds of protons ranging primarily from 2-2.7 ppm ^1H and between 30 and 40 ppm ^{13}C, similar to what has been found here for the Georgia humic acid.

In contrast to all suggestions for the source of alicyclic molecules, we conclude that alicyclic molecules containing significant carboxyl group functionality originate from hydroxyl radical oxidation reactions of lignin and lignin-rich NOM (Chen et al., 2014; Waggoner et al., 2015). The large variety of carboxyl cross peaks in this study conforms with a random assortment of structural arrangements resulting from radical-promoted polymerization reactions. Proposed pathways for the oxidation of lignin to alicyclic molecules includes side chain oxidation and demethylation followed by ring opening resulting in unsaturated aliphatic and hydroxylated carboxylic acid containing structures (Waggoner et al., 2015), a series of reactions that have been shown to occur via fungal degradation of lignin (Higuchi, 2006). Further electrocyclic polymerization with muconic acid-like structures may also lead to re-cyclization and decarboxylation through the loss of CO_2. Methyl groups, that appear to be a notable contribution to this sample, have been commonly mentioned as indicators of terpenoid sources, but we do not agree that these are exclusive indicators of terpenoids. Although these were not originally

identified per se in the structures and mechanisms proposed to be formed from lignin by Waggoner et al. (2015), these could be incorporated during polymerization reactions after demethylation or intensified relative to other groups by modifications if originally present as part of the lignin structure.

Although Waggoner et al. (2015) were unable to obtain multidimensional spectra for the hydroxyl-radical oxidized lignin sample, a study was published by Schmitt-Kopplin et al. (1998) in which a dissolved soil HA with lignin components was photoxidized and compared with the parent material using 2D NMR. These authors concluded that lignin and lipid like constituents were the most susceptible to alterations by photooxidation. Their results also support the formation of a greater variety of aliphatic proton resonances, many of which could reflect alicyclic molecules. Moreover, they observed a significant increase in methyl resonances after photo-oxidation, a hydroxyl-radical producing process.

The findings in this study are very consistent with the presence of alicyclic aliphatic compounds in humic acid. Recent work by our group (Chen et al., 2014; Waggoner et al., 2015) indicates that these alicyclic molecules can be produced abiotically from lignin via ring opening followed by electrocyclic polymerization.

CHAPTER IV

SUB-STRUCTURAL COMPONENTS OF ORGANIC COLLOIDS FROM A PU-POLLUTED SOILWITH IMPLICATIONS FOR PU MOBILIZATION

Preface

The content of this chapter was published in Environmental Science and Technology in 2017.

Below is the full citation. The formatting has been altered to incorporate the supporting

information into the body of the manuscript. See Appendix B for the copyright permission.

Reprinted (adapted) with permission from (DiDonato, N., Xu, C., Santschi, P.H., Hatcher, P.G.,

2017. Substructural Components of Organic Colloids from a Pu-Polluted Soil with Implications

for Pu Mobilization. Environmental Science & Technology. DOI: 10.1021/acs.est.6b04955).

Copyright (2017) American Chemical Society.

1. INTRODUCTION

Colloidal organic matter has been shown to mobilize reduced plutonium in contaminated soils,

sediments and surface water columns(Nelson et al., 1985; Santschi et al., 2002; Kalmykov et al.,

2010; Xu et al., 2014). Historically, primarily oxidized forms of Pu(V,VI) which are more

soluble and less easily sorbed to solid particles were of most concern for its transport (Nelson et

al., 1985). More recent studies have demonstrated the mobility of reduced forms (III, IV) due to

their high affinity for surfaces of both inorganic (Kersting et al., 1999; Novikov et al., 2006) and

organic colloids (Kalmykov et al., 2010; Santschi et al., 2002; Xu et al., 2014; Kersting, 2013).

Thus, Pu oxidation state (found primarily as IV and V) combined with environmental conditions

(Eh, pH, presence of either mobile or immobile minerals and organic matter available for

complexation or sorption/desorption reactions, etc.) largely determines the mobility of Pu in terrestrial and aquatic environments.

For example, organic rich and organic poor wetland soil from the Savanna River Site (SRS) contaminated weapons manufacturing facility was found to contain similar total Pu(IV) and Pu(V) activity, however Pu was more mobile in the organic rich soil due to formation of colloidal organic matter (1kDa-0.45 um) (Xu et al., 2014). The majority of the mobile colloidal Pu was contained in a sub-fraction that was separated via an isoelectric focusing (IEF) electrophoresis experiment and found to concentrate around the IEF range of 4.1-5.6 (defined as "SRS IEF colloid") (Xu et al., 2015). Elevated pH values, as a result of the basin closure and remediation strategies, led to Pu remobilization by complexation with colloidal organic matter and subsequent desorption from soil particles in the downstream wetland sediment. Moreover, wetland sediment Pu concentrations from this site were shown to increase with increasing nitrogen and hydroxamate content, which were proposed to originate partly from siderophores responsible for complexing Pu. Likewise, at the Rocky Flats Environmental Technology Site (RFETS), another Pu-contaminated weapons manufacturing facility, the greatest Pu-239, 240 activity was found to be associated with particulate (>0.45 um) and colloidal (≥3 kDa) organic matter fractions, rather than inorganic mineral phases(Xu et al., 2008). A Pu enriched fraction, the "RFETS IEF colloid", was sequentially isolated first by water extraction of the soil to obtain a crude colloid, and then purified using IEF electrophoresis. Elemental analysis, hydroxamate content and Pu activity of the RFETS soil, crude colloid and IEF colloid was obtained and the RFETS IEF colloid was further analyzed using a variety of techniques including mutli-dimensional NMR. The RFETS IEF colloid was found to be enriched in hydroxamate-type

compounds but additionally displayed a strong resemblance to plant cuticles based on multidimensional NMR studies (Xu et al., 2008). Hydroxamate groups linked to a cutin-like, carbohydrate containing Pu-rich organic colloid were suggested to be a source of the Pu chelation (Xu et al., 2008).

Cutin, a polyester present in the cuticle surface of many plants, is expected to accumulate in soils over time due to its hydrophobic long chain polymethylene structure resistant to microbes. It is interlinked by ester bonds, and commonly contains both primary and secondary alcohol groups (Deshmukh et al., 2003; Fang et al., 2001; Hitchcock, 1971). The biopolymer alone would not likely sequester Pu since it does not contain strong metal binding ligands and its low solubility in solution limits interaction with metal ions. There is reason to believe that strong metal binding chelates such as hydroxamate siderophores can be cross-linked to cutin polymers within soil fractions and thus, attach chelating groups to the cutin scaffold that would be excellent candidates for metal mobilizations in soil. Several studies have provided evidence for incorporation of nitrogen functional groups into organic matter (Hsu and Hatcher, 2005; McKee and Hatcher, 2010; Rillig et al., 2007), including amidation of ester groups in cutin, which can be considered a model for stable organic matter (Turner, 2007). Hydroxamate siderophores are known to interact with esters similarly (Medeiros et al., 2012). These ligands form some of the strongest complexes known (log K_{ML} = ~ 30 for Fe(III)) (Boukhalfa et al., 2007) and also chelate metals other than Fe, particularly those of similar charge to ionic radius and preference for oxygen ligands, such as the radionuclide Pu (Raymond et al., 1984). The Pu-hydroxamate stability constant is up to 5 orders of magnitude higher than for Fe (Boukhalfa et al., 2007). If a cutin biopolymer with similar structure is present in the Pu-carrying colloid extracted from soils

at the Rocky Flat's site, the possibility exists that it may have reacted with nucleophilic nitrogen species within siderophores to impart strong metal-binding characteristics.

Pu activity at both SRS and RFETS was found to be correlated with organic carbon, and even more strongly correlated with nitrogen and hydroxamate content. Nitrogen incorporated by siderpohores into cutin at the RFETS was suggested to explain some of the nitrogen correlation with Pu. However, only a small portion (~2%) of the total N content of the soil could be attributed to hydroxamate-nitrogen, so other N-containing molecules of unknown molecular composition are also likely responsible for sequestering Pu. Some of these may contain carboxyl and hydroxyl groups, phenols, salicylates, amino acids, phthalates, carbohydrates and quinone type molecules that are known to form complexes with multivalent cations (Stumm and Morgan, 2012),(Tipping, 2002). ESI-FTICRMS is a useful tool for characterizing the molecular components of organic matter due to its ultra-high resolution from which molecular formulas can be accurately assigned. This analysis was completed for the SRS IEF colloid (Xu et al., 2015), in which hydroxamate siderophore decomposition products were suggested to be present based on comparison with a hydroxamate siderophore standard. Other (CRAM-like) N-containing molecules were also identified for their potential role. However ESI-FTICRMS analysis of the RFETS IEF colloid (Xu et al., 2008) has not previously been conducted.

Thus, the goal of the current study is two-fold. First, we seek to identify additional molecular components of the RFETS IEF colloid using ESI-FTICRMS that could be responsible for Pu-chelation. For the current study we use alkali, which often provides greater (up to 80%) solvation efficiency (Stevenson, 1994), to dissolve the organic matter in the RFETS soil, water

extracted crude colloid and purified IEF colloid. While observation of the bound ligands is not possible at these low Pu concentrations, we relate the organic matter composition of each fraction to the previously reported Pu activity and chemical analysis to identify potential ligands (Xu et al., 2008). Second, we test if an amine-containing hydroxamate siderophore, desferrioxamine (DFO), may be abiotically and covalently incorporated into the biopolymer cutin via a nucleophilic addition reaction. This could explain one way hydroxamates, measured previously using spectrophotometric techniques, could persist in organic matter to chelate Pu. Our studies involve isolation of cutin from Western wheatgrass, which is the main source of organic matter to the soils at the Rocky Flats site and potentially the source of the cutin-like structures previously identified in the RFETS IEF colloid(Nelson, 2010). We attempted to bind DFO to cutin to impart metal ligation properties, attempting first to do this with well characterized tomato cutin isolates. DFO amide products have previously been synthesized under Schotten-Baumann conditions(Ihnat et al., 2000), however to our knowledge this is the first time a DFO amidation reaction has been attempted without the use of reactive acyl chlorides not expected to exist in nature.

2. MATERIALS AND METHODS

2.1 Sample preparation

Soil was collected in the summer of 2004 from the 903 'lip area' of the RFETS according to Xu et al 2008. It was dried, sieved through at 2 mm sieve and preserved under cool, dry conditions. For the original study, a crude colloid and purified IEF colloid were extracted from the soil and samples were previously analyzed using a suite of analytical tests including elemental analysis, carbohydrate, protein, metals (Fe, Al, Mn), sulfate, phosphate as well as solid state and high resolution NMR(Xu et al., 2008) and findings were published by Xu et al.(Ketterer et al., 2004;

Santschi et al., 2002; Xu et al., 2008) Briefly, the soil which contained 351±7 pCi/g of Pu-239, 240 was submerged in water to simulate runoff conditions, filtered and diafiltered to obtain a "crude colloid" (3 kDa- 0.45 μm) found to have a concentrated activity of 660±47 pCi/g of Pu, as determined by alpha spectroscopy (Ketterer et al., 2004; Santschi et al., 2002; Xu et al., 2008). The crude colloid was further purified by isoelectric focusing electrophoresis to obtain a subfraction, the "IEF colloid", with an activity of 3222±278 pCi/g Pu. For the current study, colloid samples were freshly prepared from the original preserved soil following the same procedure to obtain a newly extracted "crude colloid" and newly purified "IEF colloid" containing organic matter associated with the highest Pu activity. The colloids were freeze-dried and alkali extracts of the original soil, newly extracted crude colloid and newly purified IEF colloid were analyzed by ESI-FTICRMS.

2.2 Cutin isolation

Western wheatgrass (*Agropyron smithii*) was obtained from the re-vegetation area of RFETS and its cutin biopolymer was isolated according to a modified Deshmukh, et al (2003) procedure. In brief, plants were cleaned and de-waxed with chloroform and then treated with ammonium oxalate/oxalic acid solution (1.6% w/v / 0.4 % w/v), washed, freeze-dried and ground. Ground grass was successively extracted with chloroform, 1:1 methanol/chloroform and methanol for 12 hrs to remove soluble lipids. The material was then washed, freeze dried and treated with 4.5% sodium paraperiodate solution (adjusted to pH 4.1 with acetic acid) for 12 h to remove carbohydrates. The remaining material (largely lipid and saccharide-free) was then filtered, suspended in water, refluxed for 3 hr, washed and freeze dried. Following the above method (Deshmukh et al., 2003) established for tomato cutin isolation, it was apparent that significant amounts of lignin and cellulosic compounds remained, which are absent or not as prevalent in

tomato. Wheatgrass has been reported to contain approximately 8-15% lignin, 60-77% holocellulose and 3-10% protein depending on the season and less than 1% cutin(Kamstra et al., 1968; Soest, 1994). In order to effectively isolate cutin from wheatgrass, two steps were added to the Deshmukh et al. (2003) procedure to remove these components. First, the remaining material was further bleached with sodium hypochlorite in acetic acid to remove lignin and second, an acid hydrolysis step with 6N HCl was also performed to remove residual cellulose and proteins(Chefetz et al., 2002). Tomato cutin was isolated from organically grown tomatoes according to the original Deshmukh, et al (2003) procedure(Deshmukh et al., 2003).

2.3 Cutin + siderophore incubation

Cutin, isolated from tomato was incubated under abiotic conditions in 2 mL sterilized vials (acid washed and combusted) with 10 mg/ml of the trihydroxamate siderophore, desferrioxamine mesylate (DFOM, Sigma), which was dissolved in D_2O (Acros) and adjusted to pH=11.5 with NaOD (Cambridge), above the pKa for the terminal amine. All ingredients were sterile, including cutin, left sterile by the isolation procedure. The headspace was purged with argon to maintain anoxic conditions and vials were sealed and stored in the dark for separate 2 week and 1 month durations. After incubation, cutin was washed, shaken for 10 minutes, centrifuged and the liquid was decanted to remove any residual water soluble DFO. This process was repeated three times and the residues were freeze dried prior to analysis. The rinsate was also analyzed by multidimensional NMR for any potential soluble reaction products.

2.4 Elemental analysis

Carbon, hydrogen and nitrogen content of wheatgrass and cutin samples (wheatgrass and tomato) as well as tomato cutin after incubation with DFO were determined using a Flash 1112 Series

Elemental Analyzer. Signal response areas were calibrated to standard curves using

nicotinamide for carbon and hydrogen and aspartic acid or acetanilide for nitrogen.

2.5 Solid state NMR

Cross-polarization magic angle spinning (CPMAS ^{13}C NMR) was used to assess the success of

the wheatgrass cutin isolation procedure. Isolated cutin was ground and added to an 80 µl rotor,

spun at 10 kHz for 6400 scans using the cp.av pulse program with a recycle delay of 1 second

and a contact time of 1 ms. Spectra were externally calibrated to glycine. Direct polarization

magic angle spinning was also employed to quantify carbon signals. The hpdec.av pulse

program was used with an 18 degree pulse angle (P1=0.8 seconds) and a recycle delay of 2

seconds for 6400 scans while spinning at 10 kHz.

2.6 Multidimensional NMR

High resolution magic angle spinning (HRMAS) was also performed to obtain detailed

information concerning the connectivity of the carbon and protons within the wheatgrass and

tomato cuticle biopolymers. Approximately 20 mg of cutin was inserted into a 50 µl rotor with

approximately 30 µl DMSO. We acquired a heteronuclear single quantum coherence (^{13}C-^{1}H

HSQC) spectrum using the hsqcetgpsi2 pulse program with 90 degree pulse lengths of 3.92 µs

and 7.75 µs in the F2 (^{1}H) and F1 (^{13}C) dimensions, respectively. All spectra were obtained

using a 7 kHz spinning speed for 640 scans and a 1 ms recovery delay. The time domain for the

FID was 1024 (F2) and 200 slices were obtained in the F1 dimension. Spectra were internally

calibrated to DMSO. ^{1}H-^{1}H Total correlation spectroscopy (TOCSY) was obtained using the

mlevetgp pulse program with a mixing time of 100 ms for 128 scans and 128 slices in the F1

dimension. ^{1}H-^{1}H Correlation spectroscopy (COSY) was completed using the cosygpqf pulse

program with a 1 second delay, 128 scans and 256 slices in the F1 dimension.

2.7 ESI-FTICRMS

Soil, crude colloid and purified IEF colloid prepared for this study were stored in a freezer until use. Alkali extracts of each were prepared by dissolution in 0.01 M NaOH overnight (~12 hours) under an argon headspace. Extracts were then batch treated with a Dowex™ 50WX8-100 ion-exchange resin at a 3:1 v/v ratio and shaken for 1 hour to remove ions and prevent formation of salts during electrospray ionization. A dichloromethane rinse was performed to remove organics sorbed to the Dowex resin and these samples were analyzed separately from the resin-treated alkali extracts. Analysis was conducted in the negative ion mode on a Bruker Daltonics 12 Tesla Apex Qe FTICR-MS instrument equipped with an Apollo II ESI source and housed in the College of Sciences Major Instrumentation Cluster at Old Dominion University.   Samples were mixed 1:1 with methanol just prior to injection for a final concentration of approximately 50 ppm carbon. Spectra of 4 megawords were obtained with 300 scans, 1 millisecond source accumulation and 2 or 3 second ion accumulation, depending on the sample. Spectra were externally calibrated to polyethylene glycol and internally calibrated to a fatty acid series common to natural organic matter. Molecular formulas were assigned using an in-house Matlab code according to previously described rules and criteria (Sleighter and Hatcher, 2007; Sleighter and Hatcher, 2011). Formulas assigned from the resin-treated alkali extracts were combined with those assigned for their respective dichloromethane extract to make a combined data set for each sample. Double bond equivalents were calculated according to DBE = 1+ 0.5(2*C-H+N+P) and aromaticity indices were calculated using $A_{imod} = (1+C-0.5*O-S-0.5*H)/(C-0.5*O-S-N-P)$ for which values ≥ 0.5 and ≥ 0.67 were categorized as aromatic and condensed aromatic, respectively (Koch and Dittmar, 2006). Kendrick mass defect (for COO) was calculated according to:

$$KMD = Kendrick\ mass\ (COO) - Nominal\ Kendrick\ mass\ (COO) \tag{1}$$

$$Kendrick\ mass\ (COO) = Exact\ m/z\ value\ of\ peak * (Nominal\ mass\ COO/Exact\ mass\ COO) \tag{2}$$

3. RESULTS

3.1 Cutin isolation

Cutin was isolated from wheatgrass at the RFETS site to compare its structure with that of the RFETS colloid which was previously suggested to contain cutin-like material. We used solid state ^{13}C CPMAS NMR and multi-dimensional NMR to confirm successful isolation of the wheatgrass cutin polymer. The solid state spectrum (Figure 20b) closely resembles the spectra obtained for the more commonly studied tomato cutin (Figure 20a) (Fang et al., 2001; Deshmukh et al., 2003). The spectrum is dominated by signals from main chain aliphatic polymethylene carbons, both crystalline and amorphous at approximately 33 and 30 ppm ^{13}C respectively, as well as signals from esters and alcohols between 60 and 70 ppm ^{13}C. This is consistent with removal of lignin and cellulose as well as proteins, as supported by a 4.9% reduction in nitrogen detected by elemental analysis (Table 8). Resonances associated with lignin (56 ppm ^{13}C, methoxy; 128 and 130 ppm ^{13}C, aryl-C; and 148, 150 ppm ^{13}C, aryl-O) were successfully removed after bleaching. Cellulose (105 ppm ^{13}C, anomeric carbohydrate and 72 ppm ^{13}C, also lignin side chain) and amino acid (56 ppm ^{13}C) and were also diminished with acid hydrolysis. The primary difference between the spectra of cutin isolated from wheatgrass and that from tomato cutin is the relatively less intense signal from crystalline polymethylene units in the wheatgrass cutin.
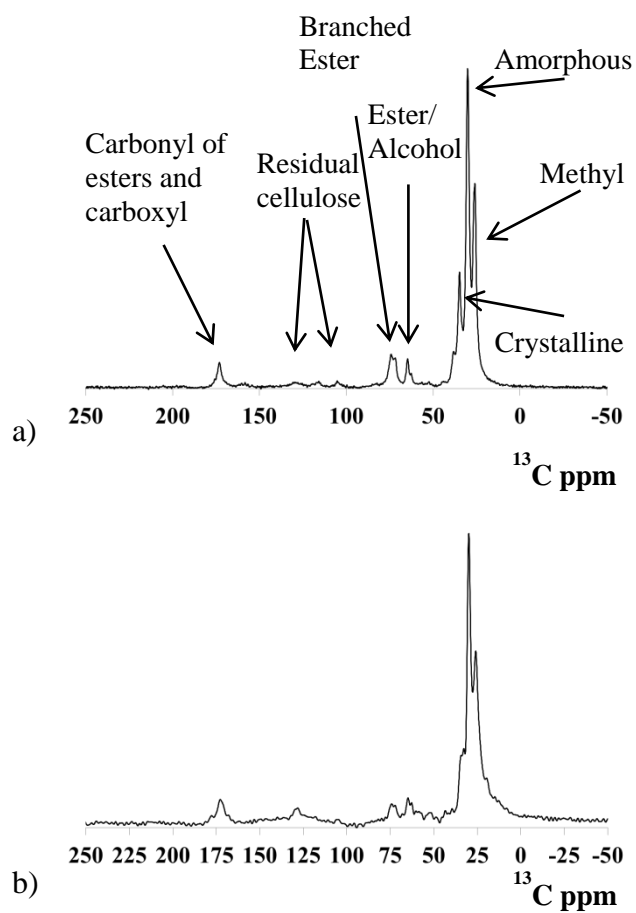
Figure 20: Solid State ^{13}C CPMAS of a) tomato cutin isolated according to Deshmukh et al.

(2003), and b) wheatgrass cutin isolated according to a modified Deshmukh et al. (2003)

Deshmukh et al. (2003) procedure.

Table 8: Wheatgrass Elemental Analysis

	Wheatgrass Prior to Cutin Isolation	Wheatgrass Cutin Post Isolation
	Replicate Masses (mg)	Replicate Masses (mg)
	0.617	0.802
	1.167	0.652
	1.06	0.578
	Elemental Analysis	Elemental Analysis
%C	45.8 +/- 0.5	41.0 +/- 1.5
%N	5.9 +/- 0.2	1.0 +/- 0.03
%H	5.7 +/- 0.1	5.6 +/- 0.5

Two-dimensional heteronuclear single quantum coherence (HSQC) NMR analysis of the wheatgrass cutin polymer which depicts resonances for only carbons that are directly bonded to protons (Figure 21) further reveals mid-chain and terminal alcohols, fatty acids, branched and straight chain esters similar to that of tomato cutin. CH_2 resonances for O-alkyl (3.9 ppm ^{1}H and 63 ppm ^{13}C) and α carbonyl-carbons (2.1 ppm ^{1}H and 34 ppm ^{13}C) of esters are shown. Methine carbons attached to α carbonyl-carbons of branched esters are also detected at 2.3 ppm ^{1}H and 42 ppm [13]C. Primary and secondary alcohols resonate at 3.3 ppm ^{1}H and 61 ppm ^{13}C, and 3.5 ppm ^{1}H and 72 ppm ^{13}C, respectively.

Total correlation spectroscopy (TOCSY), which detects only protons within the same spin system, also confirms that esters and alcohols are connected to long chain polymethylene units, as expected for cutin (Figure 22). Cross peaks appear for esters (4.0 ppm and 2.1 ppm ^{1}H) as well as alcohols (3.3 ppm ^{1}H) in the same spin system with polymethylenes (1.2 and 1.5 ppm and 1.2 and 1.3 ppm ^{1}H, respectively).Thus successful isolation of cutin from wheatgrass at the Rocky Flats Facility is confirmed by the above single and multi-dimensional NMR data.
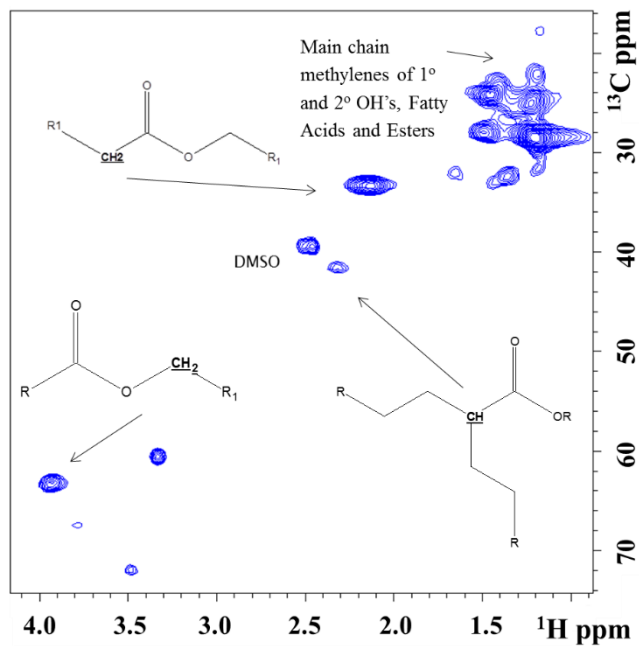
Figure 21: Wheatgrass Cutin heteronuclear single quantum coherence spectroscopy (HSQC).
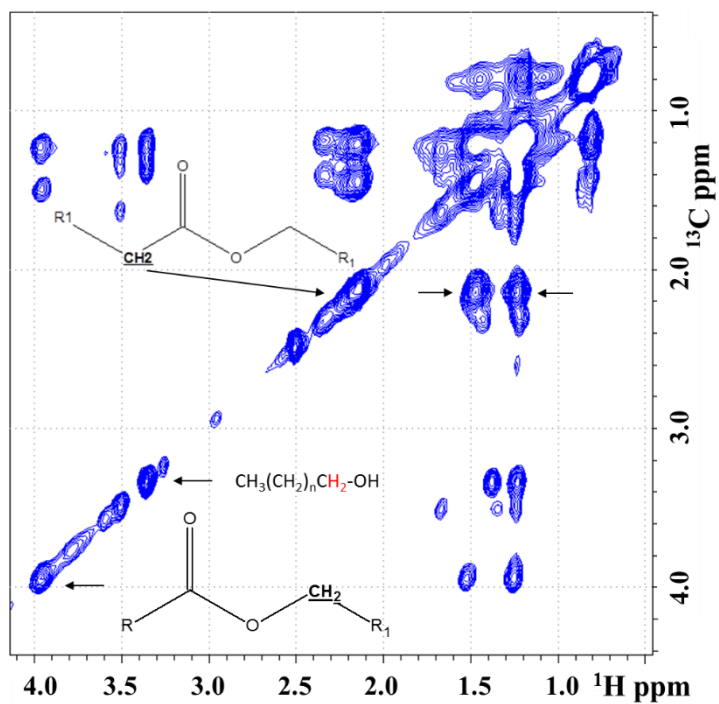
Figure 22: Wheatgrass cutin total correlation spectroscopy (TOCSY).

3.2 Wheatgrass cutin comparison with RFETS colloid structure

Figure 23 shows the HSQC spectrum for wheatgrass cutin overlaid with the spectrum from the IEF colloid. The spectra were both calibrated to DMSO and the scaling was adjusted relative to baseline to allow for detailed comparison. It is apparent from these spectra that both samples contain polymethylene CH_2 units attached to primary and secondary alcohols and fatty acids (29-30 ppm ^{13}C, 1.2-1.5 ppm ^{1}H). However the signal for O-alkyl ester resonances in the IEF colloid (blue, 66 ppm ^{13}C and 4 ppm ^{1}H) is much less intense and shifted slightly upfield from the analogously assigned peak in the spectrum for the isolated cutin (red, 63 ppm ^{13}C and 3.9 ppm ^{1}H). This could indicate some modifications to the structure in soils. This region is also shared by resonances from alcohols and ethers, generally at more upfield proton resonances (~3.3-3.6 ppm). A corresponding peak in the region where the ester α carbonyl carbon resonates is not as apparent in the spectra of the colloid, as in the grass cutin (34 ppm ^{13}C, 2.1 ppm ^{1}H).

If a nucleophilic substitution reaction occurred at the ester, the amide carbon would resonate in the region depicted by the red circle, but this is also not conspicuously observed. Further data from the TOCSY spectra of the colloid did not show similarities in the range of protons within the same spins system (mid-chain methylenes with esters, etc.) but COSY data does show similarities with the structure of the cutin isolate, particularly carbohydrates and groups α and ß to acids, esters and alcohols. Overall there is some overlap between resonances for isolated wheatgrass cutin and the colloid, and both appear to contain polymethylene units with terminal and mid-chain alcohols as well as ether, ester and potentially branched ester functionality. However, if the colloidal material originated from wheatgrass cutin at the
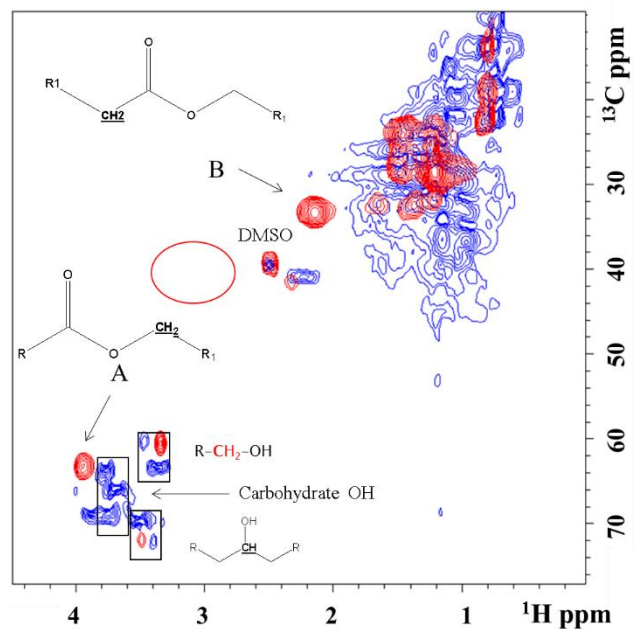
Figure 23: Overlay of heteronuclear single quantum coherence (HSQC) spectra for wheatgrass cutin (red) and RFETS soil IEF colloid (blue). Peak A is assigned to o-alkyl methylene units of the ester and peak B is assigned to methylenes attached to the carbonyl carbon of the ester. The red circle represents the location where an amide carbon would resonate upon nucleophilic substitution of the ester.

site, it is clear that some modifications, not necessarily those predicted by the proposed amidation reactions, have taken place to account for minor incongruities between the two spectra.

3.3 Cutin + siderophore incubation results

We aimed to react cutin with the hydroxamate siderophore, DFO, attempting first to do this with the well-studied and more abundant cutin biopolymer isolated from tomato. The HSQC NMR

spectrum of tomato cutin after incubation with DFO did not show any changes from that of the

cutin biopolymer prior to incubation as shown in Figure 24. As mentioned above, there is no
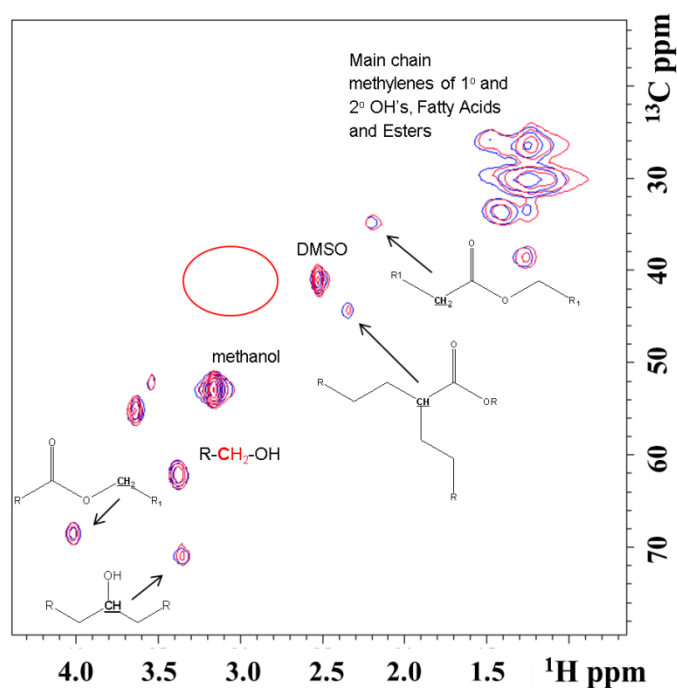
Figure 24: HSQC spectra overlay of unreacted tomato cutin (blue) and tomato cutin post incubation with DFO (red). The red circle represents the location where an amide carbon would resonate upon nucleophilic substitution of the ester.

signal in the region where the amide carbon would resonate if the amidation reaction proceeded.

Elemental analysis (Table 9) likewise, did not indicate any significant changes to suggest

incorporation of the siderophore. Therefore we cannot provide any further evidence for covalent

bonding between hydroxamates and organic matter based on this test. However,

Table 9: Elemental Analysis for Cutin-DFO Incubation

	Tomato Cutin Prior to Incubation	Tomato Cutin Post Incubation with DFO
	Replicate Masses (mg)	Replicate Masses (mg)
	1.075	1.159
	1.263	1.125
	1.293	1.096
	Elemental Analysis	Elemental Analysis
%C	67.9+/- 0.1	65.7 +/- 2.5
%N	1.4 +/- 0.1	1.5 +/- 0.1
%H	8.8 +/- 0.5	10.2 +/- 1.8

other pathways could form the proposed compounds and explain such modifications. The

method we attempted is just one, and may be hindered by other factors (i.e. physical occlusion of

active sites in the polymer, pH, etc.). Biological activity may also play an important role

inmodification of the soil aggregate and colloids, which were previously found in the dissolved

phase as small molecules.

3.4 ESI-FTICRMS

Our search for other pathways led us to employ ESI-FTICR-MS to observe molecular formulas

that might indicate other organic components of significance to the high Pu-affinity of the

colloid. Molecular formulas assigned for the base soluble organic matter in the original soil,

crude colloid and IEF colloid have been plotted in van Krevelen space in Figure 25 along with

formula type distributions. Molecular formulas for the base soluble organic matter in the soil

plot in several regions of the van Krevelen diagram, including lignin-like, carboxyl containing

aliphatic molecules, tannin-like, carbohydrate-like, aromatic and condensed aromatic.
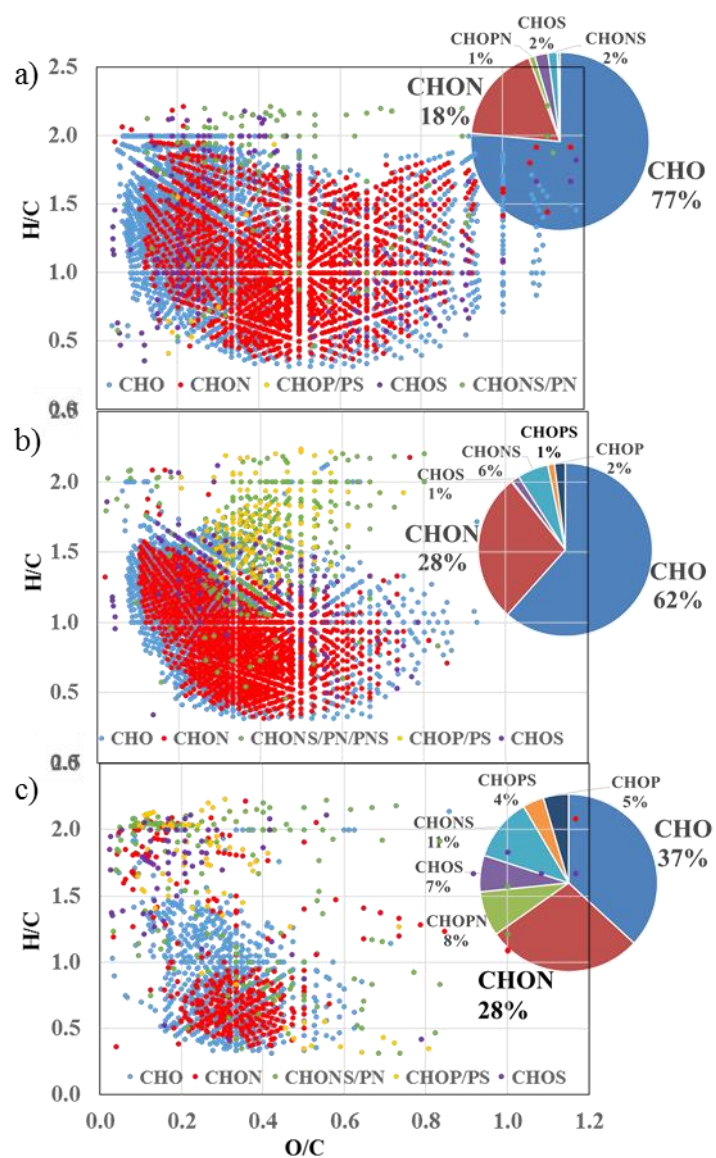
Figure 25: van Krevelen diagrams and ESI-FTICRMS assigned formula type percent weighted distributions for RFETS base extracts of a) original soil (351 ± 7 pCi Pu /g), b) crude colloid (660 pCi ± 47 Pu /g) and c) IEF colloid (3222 ± 278 pCi Pu /g). Pu activities were previously reported by Xu et al. 2008. Molecular formula percentages and weighted intensities were measured in this study.

The base soluble organic matter in the crude colloid contains molecular formulas with lower O/C ratios and relatively less molecular formulas plotting in the carbohydrate and lipid-like regions as compared to the base soluble organic matter in the soil. These compounds and high H/C aliphatic compounds are also found to a much less extent in the spectrum of the IEF colloid base soluble organic matter. It is uncertain if an aliphatic cutin or cutin-siderophore molecule could be detected using ESI-FTICRMS due to the hydrophobic nature of cutin. A portion of the polymer would need to break off, dissolve into solution and form a stable ion in the ESI source. Cutin was originally detected in a water extract, but may or may not result in an ionizable compound. Some aliphatic nitrogen containing formulas (12% of the total, mostly as CHOPN and CHONS) are detected in the alkali extract of the RFETS IEF colloid, but it is difficult to distinguish if these could represent cutin or siderophores as opposed to fatty acids typically found in organic matter. These formulas are few compared to many lower H/C nitrogen containing compounds which upon further scrutiny, show greater potential for affiliation with metals as discussed below. While there are fewer total assigned peaks for the IEF colloid extract in comparison to the whole soil and crude colloid extracts, which could be due to its lower solubility, or to lower ionization efficiency, we attribute this partly to the more homogeneous nature of this sample, being purified by the isoelectric focusing treatment (Xu et al., 2015). Although selective ionization of different classes of compounds is a consideration when using electrospray, the technique here allows us to conduct a relative comparison of heteroatom compound abundances based on relative percentages of peak numbers and intensities among the samples analyzed.

The organic matter in the crude colloid and IEF colloid originate from the soil, but were extracted from the soil with water through filtration and ultrafiltration, and the IEF colloid is a subfraction of the crude colloid after isoelectric focusing electrophoresis separation. Base extraction was not used in the original isolation, but was applied to these three samples subsequently and in parallel for their analysis. As mentioned above, the base-extracts of each subsample may represent up to 80% of the parent material from each step. 47% of the formulas in the IEF base extract were also found in the soil and crude colloid base extracts. Thus, we compare the results of the analysis to investigate trends in organic matter composition among the base extracts.

Overall, the molecular formula types for the original soil base extract were found to be predominantly CHO and CHON, with a signal intensity ratio of approximately 4:1. Together these formulas account for more than 92% of the signal intensity. This majority decreases slightly in the crude colloid and IEF colloid base extracts as the percent intensity and number of formulas (Figure 25) for other nitrogen containing formulas (CHONS, CHOPN) as well as sulfur containing formulas increase. The percent weighted intensity of all nitrogen containing formulas (CHON/NS/PN) measured in this study using ESI-FTICRMS for each organic matter extract is 21%, 34% and 47% for the whole soil, crude colloid and IEF colloid, respectively (Figure 25). This is in agreement with elevated nitrogen/C atom ratio previously measured in the IEF colloid in comparison to the soil (N/C atom ratio of 0.050 in the soil increasing to 0.203 in the IEF colloid) (Xu et al., 2008).

The relative percentages and relative weighted intensities of condensed aromatic formulas increase with increasing nitrogen and Pu concentration of the sample (soil<crude colloid<IEF colloid). Formulas with $A_{imod} \geq 0.67$ account for 9.3% of the formulas in the soil base extract (13.1% of the weighted intensity), and 19.5% (20.6% intensity) and 22% (31.2% intensity) of the molecular formulas in the crude colloid and IEF colloid base extracts, respectively. Of these, the CHO and CHON predominance still holds in all samples, with a slight increase in CHONS (<1% to 5% signal intensity, and <1% to 2% of the formulas) and CHOPN compounds (0 to 3% signal intensity and 2% of the formulas) at the expense of pure CHON composition in the IEF base extract.

A significant number of aromatic and condensed aromatic molecular formulas also contain the same Kendrick mass defect for COO. The Kendrick mass defect analysis adjusts the calculated exact mass based on the mass defect for a COO group and reveals molecular formulas for which their mass differs only by a COO. Although this technique cannot provide direct structural information, it is expected that families of these compounds with the same mass defect contain similar structures with varying numbers of carboxyl groups. Carboxyl groups were previously identified in the RFETS IEF colloid using ^{13}C CPMAS solid state NMR, but were primarily attributed to aliphatic esters and hydroxamate groups (Xu et al., 2008). This is the first time their association with aromatic and condensed aromatic molecules in the RFETS IEF colloid has been identified. Molecular formulas not included in a KMD series may still contain carboxyl groups, however these would be difficult to identify using ESI-FTICRMS data alone without the KMD series analysis. Likewise, we caution that the abundance of formulas that have the same

Kendrick mass defect does not quantitatively provide the abundance of this functional group, but only provides a relative indication of such.

Kendrick mass defect analysis reveals COO series of 2 or more (up to 4 in this case) that constitute 49% of all of the formulas (33% of the signal intensity) in the IEF colloid base extract. Although the base soluble organic matter in the soil and crude colloid samples contain more molecular formulas in a COO Kendrick mass defect series overall (up to 90% of the signal intensity and 86% of the molecular formulas) the fraction of these that are aromatic and condensed aromatic molecules increases with increasing Pu concentration of the soil, crude colloid and IEF colloid base extracts analyzed (Figure 26). 75% of the COO series formula's signal intensity in the IEF colloid base extract (73% of the COO series formulas) are aromatic

a)
b)
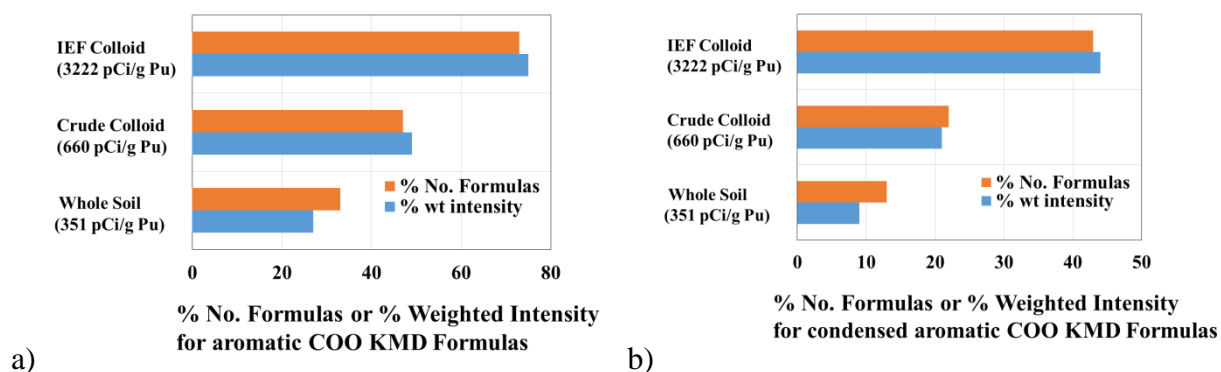
Figure 26: Percent weighted intensities (blue), and percent number of formulas (orange) for carboxyl containing aromatic (a) and carboxyl containing condensed aromatic (b) formulas of the RFETS whole soil, crude colloid and IEF colloid. Pu activities were previously reported by Xu et al. 2008. Molecular formula percentages and weighted intensities were measured in this study.

($A_{imod} \geq 0.5$) and 44% of the COO series formula's intensity are condensed aromatic ($A_{imod} \geq$ 0.67) (43% of the COO series formulas) within this sample.

Aromatic compounds that are part of a Kendrick mass defect COO series are also elevated in CHON compounds as compared to non-aromatic molecular formulas for all three samples. 19-42% and 24-50% of the signal intensity for aromatic and condensed aromatic formulas respectively are from CHON formulas (these constitute 24-55 and 31-61% of the formulas, respectively), as opposed to 0-11% of the intensity from non-aromatic formulas (0-29% of the non-aromatic formulas). All of the CHON formulas included in a KMD COO series in the IEF colloid are either aromatic or condensed aromatic (Figure 27). These formulas are mostly condensed aromatic and make up almost half of all CHON formulas (48% of CHON formulas,
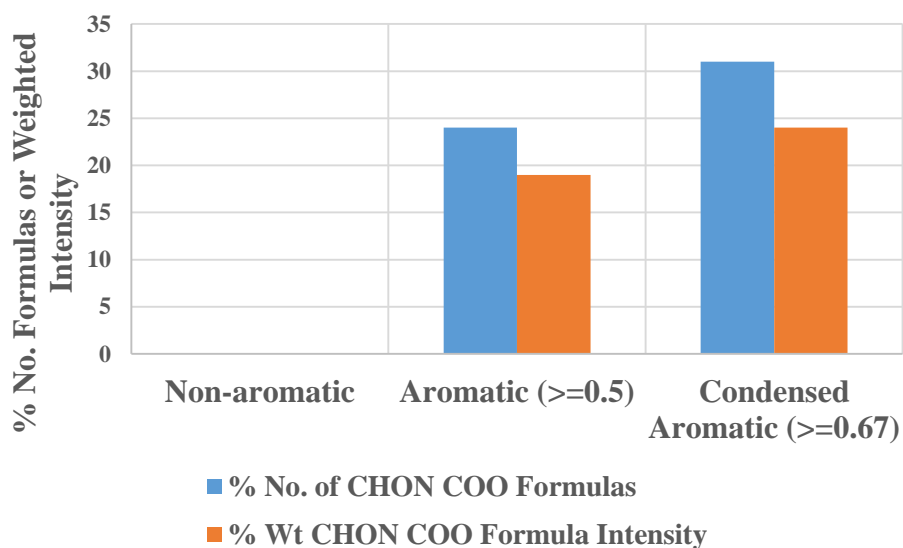
Figure 27: Percent number and weighted intensity of non-aromatic, aromatic and condensed aromatic CHON formula types that are part of a COO KMD series in the IEF colloid.

25% of CHON weighted intensity). They generally contain 1 nitrogen, 16-27 carbons, 3-10

oxygen and 10-21 double bond equivalents, with the formula $C_{23}H_{17}O_6N_1$ exemplifying one of

the most statistically representative formulas (Table 10, Figure 28).

Table 10

Statistical Composition of IEF Colloid CHON Formulas within a COO KMD Series and
representative assigned formulas

	Range	Mean	Std. Dev.	Median	Mode
#C	16-27	21.8	2.5	22	23
#H	9-23	13.9	2.9	14	15
#O	3-10	7.2	1.8	7	6
#N	1-2	1.1	0.3	1	1
O/C	0.1-0.5	0.3	0.1	0.3	0.4
DBE	10-21	16.4	2.3	16	16
A_{imod}	0.5-0.9	0.7	0.1	0.7	0.7

Formula	Exact Mass	m/z	Peak Height	DBE	A_{imod}	COO KMD Series
$C_{23}H_{17}O_6N_1$	402.098310	402.098337	1399948	16	0.66	2
$C_{22}H_{15}O_6N_1$	388.082663	388.082660	1262750	16	0.69	3
$C_{22}H_{15}O_7N_1$	404.077657	404.077575	1470616	16	0.69	3
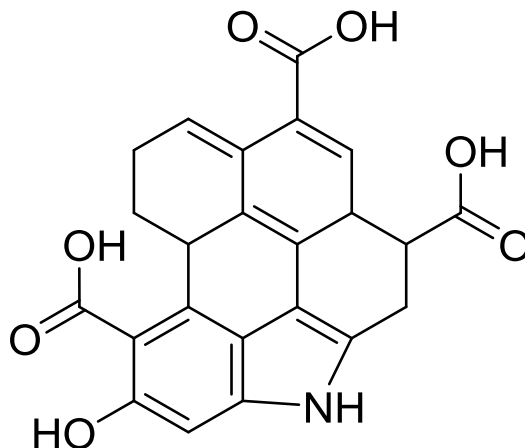
Figure 28: Hypothetical structure for a representative formula, $C_{23}H_{17}O_6N_1$ (DBE=16) detected in the IEF colloid.

4. DISCUSSION

4.1 Potential ligands

The significant percentage of CHON formulas found with COO Kendrick mass defect series led us to investigate if other carbonyl containing groups (i.e. amide, hydroxamate, etc.) may be present as well. However, KMD analysis specific for the hydroxamate group (molecules that differ in exact mass by only a CONO group) did not reveal any series and amide KMD series amounted to a total of three which plot as condensed aromatic molecules. Many of the known hydroxamate siderophores (desferrioxamine, ferrichrome, rhodotorulic acid, etc.) often contain a sufficient number of aliphatic components such that their H/C and O/C ratios plot outside of the aromatic region of the van Krevelen diagram. The region where amides and aminosugars typically plot ($1.5 \leq H/C \leq 2.0$, $0.2 \leq O/C \leq 0.7$) is notably absent of peaks compared to the remaining diagram (Sleighter and Hatcher, 2007). Although many other siderophores exist, a comparison of the ESI-FTICRMS spectra with that of a DFO standard did not provide evidence

for hydroxamates. Most of the carboxyl and nitrogen containing formulas were composed of one nitrogen, and a maximum of two, which suggests that if present, they could contain only one or two hydroxamate groups per molecule. Lack of identification may also be due either to the low solubility of these molecules in alkali, low concentration or competitive ionization. A number of other microbial and plant based siderophores that employ carboxyl amine, hydroxyl, amine or catecholate groups to chelate iron were also considered in this study. These contain a significant proportion of nitrogen and some contain an adequate number of ring structures to obtain A_{imod}>0.5 (i.e. enterobactin). However, very few, if any are expected to have condensed aromatic CHON structures identified in this study.

Hydroxamate groups were previously detected in this sample using a modified Csaky test, and while the additional information afforded in this study does not preclude the presence of hydroxamate groups and/or siderophores, including any potentially crosslinked with a cutin-polymer, it does suggest that the high Pu affinity of the colloid could be attributed to other sources as well. The relationship identified here between elevated Pu concentration and an increased proportion of condensed aromatic CHO and CHON compounds within COO homologous series, a functionality known to interact with metals, is evidence that this could be the case. Even where molecules suggestive of hydroxamate siderophores have been identified as potential Pu carriers in samples with elevated nitrogen, they represent only a small fraction of the total N content which would allow other N-containing compounds to contribute to the strong correlation between organic N and Pu concentrations (Xu et al., 2015). However, since Pu is at concentrations less than 10^{-14} M, and potential ligands are many orders of magnitude more abundant, it is very challenging to isolate the actual bound ligand(s). Unknown metal:ligand

stoichiometry, charge state and kinetics of ligand exchange vs. droplet formation in ESI-FTICRMS also limit identification of true natural bound ligands in the gas phase that would be active in solution phase (McDonald et al., 2014).

A pronounced ESI-FTICRMS signal intensity from condensed aromatic formulas was not expected for the IEF colloid, considering the solid state NMR spectra previously obtained showed relatively low (approximately 7.6%) signal attributed to aromatic carbons(Xu et al., 2008). Some of this discrepancy may be explained by the alkali extraction method. Alkali is often used in order to solubilize weakly acidic groups, and could preferentially extract more aromatic compounds if those are the primary structures in which acidic functional groups are present, as appears to be the case here. Long chain hydrophobic aliphatic polymers without acid functionality would are soluble, and more difficult to ionize in the ESI source. Additionally, NMR signal quenching due to interaction of acidic aromatic molecules with iron may have also contributed to their reduced signal in the solid state NMR. Paramagnetic species such as iron have been shown to selectively diminish signal amplitude of carbon resonances in close proximity to iron species (Pfeffer et al., 1984)·(Keeler and Maciel, 2003). Medium range effects (within a few spin systems) are usually due to shortened $T_{1\rho}$ relaxation which affects only CPMAS ^{13}C NMR spectra, however short range effects (within a few bonds) have been implicated in loss of observability for DPMAS spectra as well (Smernik and Oades, 2000). Both effects can occur within the same sample and have been noted for samples with as low as 0.02% Fe (0.00008 Fe/C mole ratio for 50% C) (Pfeffer et al., 1984). High levels of iron (2.9%) in the original whole soil from Rocky Flats initially precluded the acquisition of NMR spectra for this sample and at an order of magnitude lower concentrations in the IEF (0.1% Fe, 0.0005 Fe/C

mole ratio) iron may still selectively affect the signal of those carbons in direct contact. The

short delay time used in the DPMAS spectra may have also contributed to reduced detection of

aromatic carbons with longer relaxation times.

4.2 Implications for plutonium mobilization

Increasing percentages of N-containing and condensed aromatic formulas with increasing

purification of the organic matter (soil<crude colloid<IEF colloid) found in this study

corresponds with increasing Pu activity (351±7 pCi Pu/g soil, 660 ±47 pCi Pu/g crude colloid,

and 3222±278 pCi Pu/g IEF colloid) of the organic matter as measured previously and reported

by Xu et al (2008). Percentages of carboxyl KMD aromatic and condensed aromatic formulas

were also found to increase with purification of the organic matter. Since carboxylic acids are

known for their interaction with hard metal ions, these are one of the prime candidates for

plutonium binding in this study. Carboxyl group acidity, and thus cation exchange capacity,

increases upon attachment to condensed aromatic molecules, which also impart high surface

area. Oxidized aromatic carbon in a variety of soils, humic substances and carbon sources has

also been demonstrated to enhance sorption of metals, complexation of Fe and Pu and increase

cation exchange capacity, particularly where the density of surficial carboxylate and oxygenated

functional groups are high (Liang et al., 2006)ʼ(Wang et al., 2011) (Fujii et al., 2014) (Parsons-

Moss et al., 2014). Polycarboxylate and polyaminocarboxylic acids can form highly stable

complexes with plutonium on account of their multiple anionic oxygen donor atoms that act as

chelates.(Clark et al., 2006a) Though Pu stability constants (log K_{ML}) with polycarboxylates such

as citric acid are approximately 15 orders of magnitude lower than those for the less abundant

hydroxamates (Boukhalfa et al., 2007; Hummel et al., 2005) these and aromatic

hydroxycarboxylates such as salicylic acid are considered effective chelates (Kudo, 2001). Pu^{4+}

stability constants with N-containing nitrilotriacetate (NTA), ethylenediaminetetraacetate (EDTA), diethylenetriamine pentaacetate (DTPA) and the bidentate EDTA complex are 12.9, 26.4, 29.5 and 35.39, respectively (Clark et al., 2006a). The large coordination sphere and high charge of Pu^{4+} also allow for environmentally relevant mixed-ligand complexes such as Pu-EDTA-citrate and Pu-EDTA-carbonate, with respective stabilities of 33.95 and 35.51, comparable to Pu-hydroxamate 1:1 complexes, which are as high as 35.48 (Boukhalfa et al., 2004; Boukhalfa et al., 2007). Sorption coefficients for Pu on other N-containing compounds such as functionalized carbon nanotubes have also been found to be strong ($>10^3$) (Gupta et al., 2016; Marsh et al., 1997). Thus it is likely that carboxylate and N-functionalized polycarboxylate aromatic and condensed aromatic formulas found in this study may form strong interactions with plutonium.

The source of black-carbon like structures in soils is commonly attributed to thermally-altered (fire induced) oxidation of biomacromolecules (Hockaday et al., 2006, 2007). Lignin and cutin biopolymers from grassland vegetation are expected to be the major source of organic matter to these soils. Fires are common in grasslands and several have been documented throughout the RFETS history (Clark et al., 2006b). Oxidation of lignin by reactive oxygen species has also been suggested as a source of black carbon (Chen et al., 2014; Waggoner et al., 2015), which appears to be ubiquitous in humic acids from a variety of soils (Ikeya et al., 2015; Ohno et al., 2010; DiDonato et al., 2016). The condensed aromatic molecules we observe in this study are possibly sourced from lignin and transformed by thermogenic or radical oxidation reactions. These processes have been shown to produce condensed molecules with sufficient oxygenated functional groups such as hydroxyl and carboxyl to exhibit polarity and solubility and several structures have been proposed (Hockaday et al., 2007; Ikeya et al., 2015; Waggoner et al., 2015).

Fewer studies have identified structures for N-containing condensed aromatic compounds in soils (Wagner et al., 2015; Knicker, 2010). Pyrrole-type forms of condensed aromatics have been proposed in fire-affected soils (Wagner et al., 2015) and these structures, such as porphyrins, also form quite stable complexes with iron. The number of nitrogens observed per ion here does not support the presence of porphyrins as complete units in this sample, and the acidity of pyrrole structures alone, comparable only to water or ethanol, makes them weak ligands for metals. However, their acidity can increase when they exist as part of larger condensed structures with electron withdrawing groups such as COO groups identified here, enhancing the ability of these molecules to sequester metals. The presence of nitrogen groups overall allows for strong binding of biologically significant metals such as iron, and by way of similarity, Pu as well (Santschi et al., 1999).

Nucleophilic nitrogen species are known to be present in soils and have been shown to be incorporated into organic matter (Hsu and Hatcher, 2005; Turner, 2007). It is possible that the molecules detected in this study (see Figure 28 for a hypothetical structure) could be formed from oxidation products of lignin that undergo Diels-Alder type reactions with cis, cis-muconic acids to form alicyclic and condensed structures, as described by Waggoner et al. (2015). A Michael donor may further be incorporated via nucleophilic nitrogen species or subsequent intramolecular condensation and cyclization reactions of the incorporated nitrogen species could also occur. Alternatively, primary amines or ammonia could be incorporated by way of Paal-Knorr type reactions with 1,4 dicarbonyls under weakly acidic conditions known to form pyrrolic structures,depending on the organic matter structural precursors. Further studies to provide structural evidence and reactivity are necessary to demonstrate this.

Previous analysis of the IEF colloid using elemental analysis, spectrophotometric hydroxamate analysis and NMR, among other techniques, identified elevated nitrogen and hydroxamate content in the RFETS IEF colloid where Pu activity was found to be highest. An aliphatic cutin-siderorophore molecule was proposed to be the primary Pu chelating entity, however other molecules of unknown structure were also suggested to supplement the extreme chelating ability of siderophores and account for the stronger correlation between nitrogen and Pu concentration than could be explained solely by hydroxamates. The current study tested an abiotic reaction between cutin and the siderophore DFO, but could not provide further evidence to support a process to form this type of hydroxamate. However, our findings have revealed an additional novel sub fraction characteristic of aromatic and condensed aromatic molecules containing nitrogen moieties and carboxyl functionality using ESI-FTICRMS that had not been previously identified for this colloid. Such structures might be important in binding Pu and Fe. (Boukhalfa et al., 2007; Hummel et al., 2005). This potential Pu carrier was isolated using the same procedures as the Pu colloidal carrier identified from the SRS and is similar in size (1kDa-0.45um vs 3kDa-0.45um), isoelectric point (pH of 3.5-4.3 vs 3), and high nitrogen content. Although both colloids were found to have significant aliphatic components, the higher pH at which the Rocky Flats colloid was re-dissolved for ESI-FTICRMS analysis here may have selectively improved the solubility of more aromatic components with acid functionality. In addition, the different organic matter inputs, source vegetation and biogeochemical environments from which these colloids came may account for some of the different molecular structures observed. The intense solar radiation and large daily and seasonal temperature fluctuation of the grassland at RFETS (which was not simulated, however, in our experiments) could stimulate

natural oxidation processes that form black carbon-like structures with nitrogen incorporated, in comparison to the relatively more shaded woodland swamp of the SRS. Overall the presence of carboxyl and nitrogen species in both Pu sequestering organic matter colloids from disparate regions demonstrates the potential importance of these groups for binding metals and warrants further investigation.

CHAPTER V

SUMMARY AND CONCLUSIONS

Multi-dimensional NMR and ESI-FTICRMS have provided new insight into the structures and composition of the alkali extractable portion of soil organic matter. Evidence has surmounted that humic acids can be more distinctly defined, not merely based on procedural methods of extraction, but in terms of the predominant constituents from which they are composed. It is apparent that samples from an array of soil types and ecosystems contain three major molecular types, the relative proportion of which may indicate oxidative processes controlling their formation. Lignin-like molecules may serve as the substrate for oxidative reactions to form both condensed aromatic (black carbon-like) and alicyclic (CRAM-like) molecules. Carboxylic acid functional groups, evident from Kendrick mass defect analysis as well as multi-dimensional NMR studies, also denote oxidative transformations in the creation of alicyclic and condensed aromatic molecules.

Detailed structural configurations for alicyclic carboxylic acids have been observed for a highly aliphatic humic acid using multi-dimensional NMR studies, particularly the HSQC and HMBC experiments that can illustrate carbon-proton connectivity from one bond distance up to three, respectively. Protons of methyl groups, olefins and alcohols or other oxygenated functional groups have been detected within three bonds of carboxylic acid carbons. Further, C-H resonances match with alicyclic methylenes, while methylene's of long chains do not represent significant components, presumable due to the initial removal of lipids from this sample. This structural information together with molecular formula information obtained from ESI-

FTICRMS has allowed me to generate specific molecular structures representative of the alicyclic carboxylic acid- containing molecules detected in humic acid.

Condensed aromatic structures containing carboxylic acids and nitrogen, similarly were identified in alkali extracts of soil organic matter associated with plutonium and iron and thus proposed as potential ligands. Their origin from lignin is also possible in light of recent studies (Chen et al., 2014; Waggoner et al., 2015).

Hypothetical pathways for the formation of carboxylic acids in soils from lignin, consistent with previous studies of the photo and chemical oxidation of dissolved lignin molecules (Chen et al., 2014; Waggoner et al., 2015), have also been proposed in this dissertation. Lignin is present at an estimated 175 Gtons of carbon stored above ground (Hedges et al., 1997) and produced at approximately 12 Gtons per year (20% of terrestrial photosynthesis) (Field et al., 1998; Ruiz-Dueñas and Martínez, 2009). It is the principal carbon source which can begin to account for the vast amount of uncharacterized carbon (~915 Gtons) estimated to be stored in soils (Bianchi, 2011). Although a lack of biomarkers for lignin in sediments and soil carbon pools indicate it may not be as intractable to degradation as originally thought (Thevenot et al., 2010), evidence and a pathway for its conversion to the molecules identified and further characterized in this thesis suggest it may still contribute a considerable amount of carbon to soils in a newly identified form. Overall, a major finding of this thesis suggests that soil organic matter likely has a major source from altered lignin as a result of oxidation and the molecules produced have an affinity for Pu and Fe metals, as displayed by carboxylic acids.

REFERENCES

Adler, E., 1977. Lignin chemistry—past, present and future. Wood Science and Technology 11, 169-218.

Antony, R., Grannas, A.M., Willoughby, A.S., Sleighter, R.L., Thamban, M., Hatcher, P.G., 2014. Origin and Sources of Dissolved Organic Matter in Snow on the East Antarctic Ice Sheet. Environmental Science & Technology 48, 6151-6159.

Argyropoulos, D.S., Jurasek, L., Krištofová, L., Xia, Z., Sun, Y., Paluš, E., 2002. Abundance and reactivity of dibenzodioxocins in softwood lignin. Journal of Agricultural and Food Chemistry 50, 658-666.

Bianchi, T.S., 2011. The role of terrestrially derived organic carbon in the coastal ocean: A changing paradigm and the priming effect. Proceedings of the National Academy of Sciences 108, 19473-19481.

Boerjan, W., Ralph, J., Baucher, M., 2003. Lignin biosynthesis. Annual Review of Plant Biology 54, 519-546.

Boukhalfa, H., Reilly, S.D., Neu, M.P., 2007. Complexation of Pu(IV) with the Natural Siderophore Desferrioxamine B and the Redox Properties of Pu(IV)(siderophore) Complexes. Inorganic Chemistry 46, 1018-1026.

Boukhalfa, H., Reilly, S.D., Smith, W.H., Neu, M.P., 2004. EDTA and Mixed-Ligand Complexes of Tetravalent and Trivalent Plutonium. Inorganic Chemistry 43, 5816-5823.

Carlile, M.J., Watkinson, S.C., Gooday, G.W., 2001. The Fungi, Second ed. Elsevier Ltd, Hungary.

Chefetz, B., Chen, Y., Clapp, C.E., Hatcher, P.G., 2000. Characterization of organic matter in soils by thermochemolysis using tetramethylammonium hydroxide (TMAH). Soil Science Society of America Journal 64, 583-589.

Chefetz, B., Salloum, M.J., Deshmukh, A.P., Hatcher, P.G., 2002. Structural Components of Humic Acids as Determined by Chemical Modifications and Carbon-13 NMR, Pyrolysis-, and

Thermochemolysis-Gas Chromatography/Mass Spectrometry. Soil Science Society of America Journal 66, 1159-1171.

Chen, C.-L., Chang, H.-M., Kirk, T.K., 1983. Carboxylic Acids Produced Through Oxidative Cleavage of Aromatic Rings During Degradation of Lignin in Spruce Wood by Phanerochaete Chrysosporium. Journal of Wood Chemistry and Technology 3, 35-57.

Chen, H., Abdulla, H.A.N., Sanders, R.L., Myneni, S.C.B., Mopper, K., Hatcher, P.G., 2014. Production of black carbon-like and aliphatic molecules from terrestrial dissolved organic matter in the presence of sunlight and iron. Environmental Science & Technology Letters 1, 399-404.

Chen, Y., Schnitzer, M., 1978. The surface tension of Aqueous Solutions of Soil Humic Substances. Soil Science 125, 7-15.

Claridge, T.D.W., 2009. High-resolution NMR Techniques in Organic Chemistry. Elsevier, Oxford.

Clark, D.L., Hecker, S.S., Jarvinen, G.D., Neu, M.P., 2006a. Plutonium, in: Morss, L.R., Edelstein, N.M., Fuger, J. (Eds.), The Chemistry of the Actinide and Transactinide Elements. Springer Netherlands, Dordrecht, pp. 813-1264.

Clark, D.L., Janecky, D.R., Lane, L.J., 2006b. Science-based cleanup of Rocky Flats. Physics Today 59, 34-40.

Cook, R.L., McIntyre, D.D., Langford, C.H., Vogel, H.J., 2003. A comprehensive liquid-state heteronuclear and multidimensional NMR study of Laurentian fulvic acid. Environmental Science & Technology 37, 3935-3944.

Deshmukh, A.P., Pacheco, C., Hay, M.B., Myneni, S.C.B., 2007. Structural environments of carboxyl groups in natural organic molecules from terrestrial systems. Part 2: 2D NMR spectroscopy. Geochimica et Cosmochimica Acta 71, 3533-3544.

Deshmukh, A.P., Simpson, A.J., Hatcher, P.G., 2003. Evidence for cross-linking in tomato cutin using HR-MAS NMR spectroscopy. Phytochemistry 64, 1163-1170.

Devlin, H.R., Harris, I.J., 1984. Mechanism of the oxidation of aqueous phenol with dissolved oxygen. Industrial & Engineering Chemistry Fundamentals 23, 387-392.

DiDonato, N., Chen, H., Waggoner, D., Hatcher, P.G., 2016. Potential origin and formation for molecular components of humic acids in soils. Geochimica et Cosmochimica Acta 178, 210-222.

U.S. Department of Energy, 2013. Site Operations Guide Appendix G Wildland Fire Management Plan for the Rocky Flats, Colorado, Site www.lm.doe.gov/Rocky_Flats/SOG/appxg.pdf, 13.

Fang, X., Qiu, F., Yan, B., Wang, H., Mort, A.J., Stark, R.E., 2001. NMR studies of molecular structure in fruit cuticle polyesters. Phytochemistry 57, 1035-1042.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., Falkowski, P., 1998. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. Science 281, 237-240.

Flaig, W., 1964. Effects of micro-organisms on the transformation of lignin to humic substances Geochimica et Cosmochimica Acta 28, 1523-1535.

Fujii, M., Imaoka, A., Yoshimura, C., Waite, T.D., 2014. Effects of Molecular Composition of Natural Organic Matter on Ferric Iron Complexation at Circumneutral pH. Environmental Science & Technology 48, 4414-4424.

Gautier-Luneau, I., Merle, C., Phanon, D., Lebrun, C., Biaso, F., Serratrice, G., Pierre, J.L., 2005. New trends in the chemistry of iron(III) citrate complexes: Correlations between X-ray structures and solution species probed by electrospray mass spectrometry and kinetics of iron uptake from citrate by iron chelators. Chemistry-a European Journal 11, 2207-2219.

Grinhut, T., Hertkorn, N., Schmitt-Kopplin, P., Hadar, Y., Chen, Y., 2011. Mechanisms of Humic Acids Degradation by White Rot Fungi Explored Using 1H NMR Spectroscopy and FTICR Mass Spectrometry. Environmental Science & Technology 45, 2748-2754.

Guggenberger, G., 2005. Humification and Mineralization in Soils, in: Varma, A., Buscot, F. (Eds.), Microorganisms in Soils: Roles in Genesis and Functions. Springer Berlin, pp. 85-106.

Gupta, N.K., Sengupta, A., Boda, A., Adya, V.C., Ali, S.M., 2016. Oxidation state selective sorption behavior of plutonium using N,N-dialkylamide functionalized carbon nanotubes: experimental study and DFT calculation. RSC Advances 6, 78692-78701.

Gupta, N.S., 2014. Biopolymers: A molecular paleontology approach. Springer Netherlands.

Haider, K., Martin, J.P., 1967. Synthesis and Transformation of Phenolic Compounds by Epicoccum nigrum in Relation to Humic Acid Formation1. Soil Science Society of America Journal 31, 766-772.

Hartman, B.E., Chen, H., Hatcher, P.G., 2015. A non-thermogenic source of black carbon in peat and coal. International Journal of Coal Geology 144–145, 15-22.

Hatcher, P.G., 1988. Dipolar-dephasing 13C NMR studies of decomposed wood and coalified xylem tissue: evidence for chemical structural changes associated with defunctionalization of lignin structural units during coalification. Energy & Fuels 2, 48-58.

Hatcher, P.G., Clifford, D.J., 1994. Flash Pyrolysis and in-situ methylation of humic acids from soil. Organic Geochemistry 21, 1081-1092.

Hayes, M.H.B., Wilson, W.S., 1997. Humic Substances, Peats and Sludges: Health and Environmental Aspects. Elsevier Science.

Hedges, J.I., Keil, R.G., Benner, R., 1997. What happens to terrestrial organic matter in the ocean? Organic Geochemistry 27, 195-212.

Heitner, C., Dimmel, D., Schmidt, J., 2016. Lignin and Lignans: Advances in Chemistry. CRC Press.

Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., Kettrup, A., Hedges, J.I., 2006. Characterization of a major refractory component of marine dissolved organic matter. Geochimica et Cosmochimica Acta 70, 2990-3010.

Hertkorn, N., Harir, M., Koch, B.P., Michalke, B., Schmitt-Kopplin, P., 2013. High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. Biogeosciences 10, 1583-1624.

Hertkorn, N., Permin, A., Perminova, I., Kovalevskii, D., Yudov, M., Petrosyan, V., Kettrup, A., 2002. Comparative analysis of partial structures of a peat humic and fulvic acid using one- and two-dimensional nuclear magnetic resonance spectroscopy. Journal of Environmental Quality 31, 375-387.

Higuchi, T., 2004. Microbial degradation of lignin: Role of lignin peroxidase, manganese peroxidase, and laccase. Proceedings of the Japan Academy Series B-Physical and Biological Sciences 80, 204-214.

Hitchcock, C., Nichols, B.W., 1971. Plant Lipid Biochemistry. Academic Press, London.

Hockaday, W.C., Grannas, A.M., Kim, S., Hatcher, P.G., 2006. Direct molecular evidence for the degradation and mobility of black carbon in soils from ultrahigh-resolution mass spectral analysis of dissolved organic matter from a fire-impacted forest soil. Organic Geochemistry 37, 501-510.

Hockaday, W.C., Grannas, A.M., Kim, S., Hatcher, P.G., 2007. The transformation and mobility of charcoal in a fire-impacted watershed. Geochimica et Cosmochimica Acta 71, 3432-3445.

Hsu, P.-H., Hatcher, P.G., 2005. New evidence for covalent coupling of peptides to humic acids based on 2D NMR spectroscopy: A means for preservation. Geochimica et Cosmochimica Acta 69, 4521-4533.

Hummel, W., Anderegg, G., Rao, L., Puigdomenech, I., Tochiyama, O., 2005. Chemical Thermodynamics of Compounds and Complexes of U, Np, Pu, Am, Tc, Se, Ni, and Zr with Selected Organic Ligands OECD-NEA, Elsevier, Boston.

Hur, M., Yeo, I., Park, E., Kim, Y.H., Yoo, J., Kim, E., No, M.-h., Koh, J., Kim, S., 2010. Combination of Statistical Methods and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for More Comprehensive, Molecular-Level Interpretations of Petroleum Samples. Analytical Chemistry 82, 211-218.

Ihnat, P.M., Vennerstrom, J.L., Robinson, D.H., 2000. Synthesis and solution properties of deferoxamine amides. Journal of Pharmaceutical Sciences 89, 1525-1536.

Ikeya, K., Sleighter, R.L., Hatcher, P.G., Watanabe, A., 2013. Fourier transform ion cyclotron resonance mass spectrometric analysis of the green fraction of soil humic acids. Rapid Communications in Mass Spectrometry 27, 2559-2568.

Ikeya, K., Sleighter, R.L., Hatcher, P.G., Watanabe, A., 2015. Characterization of the chemical composition of soil humic acids using Fourier transform ion cyclotron resonance mass spectrometry. Geochimica et Cosmochimica Acta 153, 169-182.

Kalmykov, S.N., Batuk, O.N., Bouby, M., Denecke, M.A., Novikov, A.P., Perminova, I.V., Shcherbina, N.S., 2010. Plutonium speciation and formation of nanoparticles with natural organic matter in contaminated environment. Geochimica et Cosmochimica Acta 74, A490-A490.

Kamga, A.W., Behar, F., Hatcher, P.G., 2014. Quantitative Analysis of Long Chain Fatty Acids Present in a Type I Kerogen Using Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Compared with BF3/MeOH Methylation/GC-FID. Journal of the American Society for Mass Spectrometry 25, 880-890.

Kamstra, L.D., Schentzel, D.L., Lewis, J.K., Elderkin, R.L., 1968. Maturity Studies with Western Wheatgrass. Journal of Range Management 21, 235-239.

Keeler, C., Maciel, G.E., 2003. Quantitation in the solid-state C-13 NMR analysis of soil and organic soil fractions. Analytical Chemistry 75, 2421-2432.

Kelleher, B.P., Simpson, A.J., 2006. Humic substances in soils: Are they really chemically distinct? Environmental Science & Technology 40, 4605-4611.

Kelleher, B.P., Simpson, M.J., Simpson, A.J., 2006. Assessing the fate and transformation of plant residues in the terrestrial environment using HR-MAS NMR spectroscopy. Geochimica et Cosmochimica Acta 70, 4080-4094.

Kersting, A.B., 2013. Plutonium Transport in the Environment. Inorganic Chemistry 52, 3533-3546.

Kersting, A.B., Efurd, D.W., Finnegan, D.L., Rokop, D.J., Smith, D.K., Thompson, J.L., 1999. Migration of plutonium in ground water at the Nevada Test Site. Nature 397, 56-59.

Ketterer, M.E., Hafer, K.M., Link, C.L., Kolwaite, D., Wilson, J., Mietelski, J.W., 2004. Resolving global versus local/regional Pu sources in the environment using sector ICP-MS. Journal of Analytical Atomic Spectrometry 19, 241-245.

Kim, S., Kramer, R.W., Hatcher, P.G., 2003. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. Analytical Chemistry 75, 5336-5344.

Kingery, W.L., Simpson, A.J., Hayes, M.H.B., Locke, M.A., Hicks, R.P., 2000. The application of multidimensional NMR to the study of soil humic substances. Soil Science 165, 483-494.

Knicker, H., 2010. "Black nitrogen" – an important fraction in determining the recalcitrance of charcoal. Organic Geochemistry 41, 947-950.

Koch, B.P., 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, 926-932.

Koch, B.P., Ludwichowski, K.-U., Kattner, G., Dittmar, T., Witt, M., 2008. Advanced characterization of marine dissolved organic matter by combining reversed-phase liquid chromatography and FT-ICR-MS. Marine Chemistry 111, 233-241.

Koenig, A.B., Sleighter, R.L., Salmon, E., Hatcher, P.G., 2010. NMR structural characterization of *Quercus alba* (White Oak) degraded by the brown rot fungus, *Laetiporus sulphureus*. Journal of Wood Chemistry and Technology 30, 61-85.

Kögel-Knabner, I., Zech, W., Hatcher, P.G., 1991. Chemical Structural Studies of Forest Soil Humic Acids: Aromatic Carbon Fraction. Soil Science Society of America Journal 55, 241-247.

Kramer, R.W., Kujawinski, E.B., Hatcher, P.G., 2004. Identification of Black Carbon Derived Structures in a Volcanic Ash Soil Humic Acid by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Environmental Science & Technology 38, 3387-3395.

Kramer, R.W., Kujawinski, E.B., Zang, X., Green-Church, K.B., Jones, R.B., Freitas, M.A., Hatcher, P.G., 2001. Studies of the structure of humic substances by electrospray ionization coupled to a quadrupole-time of flight (QQ-TOF) mass spectrometer, in: Ghabbour, E.A., Davies, G. (Eds.), Humic Substances: Structures, Models and Functions. The Royal Society of Chemistry, Cambridge, pp. 95-107.

Kudo, A., 2001. Plutonium in the Environment. Elsevier Science.

Kumada, K., 1987. Chemistry of Soil Organic Matter. Elsevier, Amsterdam.

Kurbatov, I.M., 1968. The question of the genesis of peat and its humic acids., in: Robertson, R.A. (Ed.), Transactions of the 2nd International Peat Congress, Leningrad, HMSO, Edinburgh., pp. 133-137.

Lam, B., Baer, A., Alaee, M., Lefebvre, B., Moser, A., Williams, A., Simpson, A.J., 2007. Major structural components in freshwater dissolved organic matter. Environmental Science & Technology 41, 8240-8247.

Lechtenfeld, O.J., Hertkorn, N., Shen, Y., Witt, M., Benner, R., 2015. Marine sequestration of carbon in bacterial metabolites. Nature Communications 6, 8.

Leonowicz, A., Cho, N.S., Luterek, J., Wilkolazka, A., Wojtas-Wasilewska, M., Matuszewska, A., Hofrichter, M., Wesenberg, D., Rogalski, J., 2001. Fungal laccase: properties and activity on lignin. Journal of Basic Microbiology 41, 185-227.

Liang, B., Lehmann, J., Solomon, D., Kinyangi, J., Grossman, J., O'Neill, B., Skjemstad, J.O., Thies, J., Luizao, F.J., Petersen, J., Neves, E.G., 2006. Black Carbon increases cation exchange capacity in soils. Soil Science Society of America Journal 70, 1719-1730.

Mao, J., Chen, N., Cao, X., 2011. Characterization of humic substances by advanced solid state NMR spectroscopy: Demonstration of a systematic approach. Organic Geochemistry 42, 891-902.

Marsh, S.F., Jarvinen, G.D., Kim, J.S., Nam, J., Bartsch, R.A., 1997. New bifunctional anion-exchange resins for nuclear waste treatment. Reactive and Functional Polymers 35, 75-80.

McDonald, L.W., Campbell, J.A., Clark, S.B., 2014. Failure of ESI Spectra to Represent Metal-Complex Solution Composition: A Study of Lanthanide–Carboxylate Complexes. Analytical Chemistry 86, 1023-1029.

McKee, G.A., Hatcher, P.G., 2010. Alkyl amides in two organic-rich anoxic sediments: A possible new abiotic route for N sequestration. Geochimica et Cosmochimica Acta 74, 6436-6450.

Medeiros, M., Orth, E.S., Manfredi, A.M., Pavez, P., Micke, G.A., Kirby, A.J., Nome, F., 2012. Dephosphorylation Reactions of Mono-, Di-, and Triesters of 2,4-Dinitrophenyl Phosphate with Deferoxamine and Benzohydroxamic Acid. Journal of Organic Chemistry 77, 10907-10913.

Mesfioui, R., Love, N.G., Bronk, D.A., Mulholland, M.R., Hatcher, P.G., 2012. Reactivity and chemical characterization of effluent organic nitrogen from wastewater treatment plants determined by Fourier transform ion cyclotron resonance mass spectrometry. Water Research 46, 622-634.

Nebbioso, A., Piccolo, A., 2013. Molecular characterization of dissolved organic matter (DOM): a critical review. Analytical and Bioanalytical Chemistry 405, 109-124.

Nebbioso, A., Piccolo, A., Lamshoft, M., Spiteller, M., 2014. Molecular characterization of an end-residue of humeomics applied to a soil humic acid. RSC Advances 4, 23658-23665.

Nelson, D.M., Penrose, W.R., Karttunen, J.O., Mehlhaff, P., 1985. Effects of dissolved organic carbon on the adsorption properties of plutonium in natural waters. Environmental Science & Technology 19, 127-131.

Nelson, J.K., 2010. Vascular Flora of the Rocky Flats Area, Jefferson County, Colorado, USA Phytologia 92, 121-150.

Novikov, A.P., Kalmykov, S.N., Utsunomiya, S., Ewing, R.C., Horreard, F., Merkulov, A., Clark, S.B., Tkachev, V.V., Myasoedov, B.F., 2006. Colloid Transport of Plutonium in the Far-Field of the Mayak Production Association, Russia. Science 314, 638-641.

Ohno, T., He, Z.Q., Sleighter, R.L., Honeycutt, C.W., Hatcher, P.G., 2010. Ultrahigh resolution mass spectrometry and indicator species analysis to identify marker components of soil- and plant biomass-derived organic matter fractions. Environmental Science & Technology 44, 8594-8600.

Olk, C., Dancel, M.C., Moscoso, E., Jimenez, R.R., Dayrit, F.M., 2002. Accumulation of lignin residues in organic matter fractions of lowland rice soils; A pyrolysis-GC-MS study. Soil Science 167, 590-606.

Page, S.E., Kling, G.W., Sander, M., Harrold, K.H., Logan, J.R., McNeill, K., Cory, R.M., 2013. Dark Formation of Hydroxyl Radical in Arctic Soil and Surface Waters. Environmental Science & Technology 47, 12860-12867.

Page, S.E., Sander, M., Arnold, W.A., McNeill, K., 2012. Hydroxyl Radical Formation upon Oxidation of Reduced Humic Acids by Oxygen in the Dark. Environmental Science & Technology 46, 1590-1597.

Parsons-Moss, T., Tüysüz, H., Wang, D., Jones, S., Olive, D., Nitsche, H., 2014. Plutonium sorption to nanocast mesoporous carbon. Radiochimica Acta 102, 489.

Perdue, E.M., Hertkorn, N., Kettrup, A., 2007. Substitution patterns in aromatic rings by increment analysis. Model development and application to natural organic matter. Analytical Chemistry 79, 1010-1021.

Pfeffer, P.E., Gerasimowicz, W.V., Piotrowski, E.G., 1984. Effect of paramagnetic iron on quantitation in carbon-13 cross polarization magic angle spinning nuclear magnetic resonance spectrometry of heterogeneous environmental matrixes. Analytical Chemistry 56, 734-741.

Raymond, K.N., Muller, G., Matzanke, B.F., 1984. Complexation of iron by siderophores- a review of their solution and structural chemistry and biological function. Topics in Current Chemistry 123, 49-102.

Reemtsma, T., 2009. Determination of molecular formulas of natural organic matter molecules by (ultra-) high-resolution mass spectrometry: Status and needs. Journal of Chromatography A 1216, 3687-3701.

Repeta, D.J., 2015. Chapter 2 - Chemical Characterization and Cycling of Dissolved Organic Matter A2 - Hansell, Dennis A, in: Carlson, C.A. (Ed.), Biogeochemistry of Marine Dissolved Organic Matter (Second Edition). Academic Press, Boston, pp. 21-63.

Riederer, M., Muller, C., 2008. Annual Plant Reviews, Biology of the Plant Cuticle. Wiley.

Rillig, M.C., Caldwell, B.A., Wosten, H.A.B., Sollins, P., 2007. Role of proteins in soil carbon and nitrogen storage: controls on persistence. Biogeochemistry 85, 25-44.

Ruiz-Dueñas, F.J., Martínez, Á.T., 2009. Microbial degradation of lignin: how a bulky recalcitrant polymer is efficiently recycled in nature and how we can take advantage of this. Microbial biotechnology 2, 164-177.

Santschi, P.H., Guo, L., Means, J.C., Ravischandran, M., 1999. Natural Organic Matter Binding of Trace Metals and Trace Organic Contaminants in Estuaries, in: Bianchi, T.S., Pennock, J.R., Twilley, R.R. (Eds.), Biogeochemistry of Gulf of Mexico Estuaries. Wiley, New York.

Santschi, P.H., Roberts, K.A., Guo, L.D., 2002. Organic nature of colloidal actinides transported in surface water environments. Environmental Science & Technology 36, 3711-3719.

Schmitt-Kopplin, P., Hertkorn, N., Schulten, H.-R., Kettrup, A., 1998. Structural changes in a dissolved soil humic acid during photochemical degradation processes under O2 and N2 atmosphere. Environmental Science & Technology 32, 2531-2541.

Schnitzer, M., Khan, S.U., 1972. Humic substances in the environment. Marcel Dekker, Inc., New York.

Schulten, H.R., Schnitzer, M., 1992. Structural studies on soil humic acids by curie-point pyrolysis-gas chromatography mass-spectrometry Soil Science 153, 205-224.

Shende, R.V., Levec, J., 2000. Subcritical aqueous-phase oxidation kinetics of acrylic, maleic, fumaric, and muconic Acids. Industrial & Engineering Chemistry Research 39, 40-47.

Simpson, A.J., Burdon, J., Graham, C.L., Hayes, M.H.B., Spencer, N., Kingery, W.L., 2001. Interpretation of heteronuclear and multidimensional NMR spectroscopy of humic substances. European Journal of Soil Science 52, 495-509.

Simpson, A.J., Kingery, W.L., Hatcher, P.G., 2003. The identification of plant derived structures in humic materials using three-dimensional NMR spectroscopy. Environmental Science & Technology 37, 337-342.

Simpson, A.J., McNally, D.J., Simpson, M.J., 2011. NMR spectroscopy in environmental research: From molecular interactions to global processes. Progress in Nuclear Magnetic Resonance Spectroscopy 58, 97-175.

Singer, G.A., Fasching, C., Wilhelm, L., Niggemann, J., Steier, P., Dittmar, T., Battin, T.J., 2012. Biogeochemically diverse organic matter in Alpine glaciers and its downstream fate. Nature Geosci 5, 710-714.

Sleighter, R.L., Chen, H., Wozniak, A.S., Willoughby, A.S., Caricasole, P., Hatcher, P.G., 2012. Establishing a Measure of Reproducibility of Ultrahigh-Resolution Mass Spectra for Complex Mixtures of Natural Organic Matter. Analytical Chemistry 84, 9184-9191.

Sleighter, R.L., Hatcher, P.G., 2007. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. Journal of Mass Spectrometry 42, 559-574.

Sleighter, R.L., Hatcher, P.G., 2008. Molecular characterization of dissolved organic matter (DOM) along a river to ocean transect of the lower Chesapeake Bay by ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Marine Chemistry 110, 140-152.

Sleighter, R.L., Hatcher, P.G., 2011. Fourier transform mass spectrometry for the molecular level characterization of natural organic matter: Instrument capabilities, applications, and limitations, in: Nikolic, G. (Ed.), Fourier Transforms - Approach to Scientific Principles. INTECH Open Access Publisher, p. 482.

Sleighter, R.L., McKee, G.A., Hatcher, P.G., 2009. Direct Fourier transform mass spectral analysis of natural waters with low dissolved organic matter. Organic Geochemistry 40, 119-125.

Sleighter, R.L., McKee, G.A., Liu, Z., Hatcher, P.G., 2008. Naturally present fatty acids as internal calibrants for Fourier transform mass spectra of dissolved organic matter. Limnology and Oceanography: Methods 6, 246-253.

Smernik, R.J., Oades, J.M., 2000. The use of spin counting for determining quantitation in solid state 13C NMR spectra of natural organic matter: 1. Model systems and the effects of paramagnetic impurities. Geoderma 96, 101-129.

Soest, P.J.V., 1994. Nutritional Ecology of the Ruminant, 2nd Edition ed. Cornell University Press, Ithica, New York.

Sollins, P., Swanston, C., Kramer, M., 2007. Stabilization and destabilization of soil organic matter—a new focus. Biogeochemistry 85, 1-7.

Sparks, L.D., 2003. Environmental Soil Chemistry, Second ed. Elsevier Science.

Stevenson, F.J., 1994. Humus Chemistry: Genesis, Composition, Reactions. John Wiley & Sons, Canada.

Stubbins, A., Spencer, R.G.M., Chen, H.M., Hatcher, P.G., Mopper, K., Hernes, P.J., Mwamba, V.L., Mangangu, A.M., Wabakanghanzi, J.N., Six, J., 2010. Illuminated darkness: Molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry. Limnology and Oceanography 55, 1467-1477.

Stumm, W., Morgan, J.J., 2012. Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters. Wiley, New York.

Sutton, R., Sposito, G., 2005. Molecular structure in soil humic substances: The new view. Environmental Science & Technology 39, 9009-9015.

Thevenot, M., Dignac, M.F., Rumpel, C., 2010. Fate of lignins in soils: A review. Soil Biology & Biochemistry 42, 1200-1211.

Tipping, E., 2002. Cation Binding by Humic Substances. Cambridge University Press, Cambridge, England.

Turner, J.W., 2007. High Resolution Nuclear Magnetic Resonance Investigations of Polymethylenic Plant Biopolymers: Structural Determinations and Post-depositional Ammonia Nitrogen Incorporation, Chemistry. Ohio State University, Columbus, OH, p. 190.

URL: International Humic Substances Society Natural Organic Matter Research. http://www.humicsubstances.org/sugar.html. Accessed: November 22, 2016. Source: Laboratory of Chemistry of the Colloids of Soils and Waters, The University of Birmingham, Edgbaston, England.

URL: Official Soil Series Descriptions (OSD) with series extent mapping capabilities. https://soilseries.sc.egov.usda.gov/. Accessed: September, 2015.

URL: USGS Federal Fire Occurrence Website. [http://wildfire.cr.usgs.gov/firehistory/index.html]. Accessed: August 2016.

van Krevelen, D., 1950. Graphical-statistical method for the study of structure and reaction processes of coal Fuel 29, 269-284.

Waggoner, D.C., Chen, H., Willoughby, A.S., Hatcher, P.G., 2015. Formation of black carbon-like and alicyclic aliphatic compounds by hydroxyl radical initiated degradation of lignin. Organic Geochemistry 82, 69-76.

Wagner, S., Dittmar, T., Jaffe, R., 2015. Molecular characterization of dissolved black nitrogen via electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Organic Geochemistry 79, 21-30.

Wang, X.S., Miao, H.H., He, W., Shen, H.L., 2011. Competitive adsorption of Pb (II), Cu (II), and Cd (II) ions on wheat-residue derived black carbon. Journal of Chemical and Engineering Data 56, 444-449.

Waska, H., Koschinsky, A., Ruiz Chancho, M.J., Dittmar, T., 2015. Investigating the potential of solid-phase extraction and Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) for the isolation and identification of dissolved metal–organic complexes from natural waters. Marine Chemistry 173, 78-92.

Watanabe, A., Fujimori, H., Nagai, Y., Miyajima, T., Kuwatsuka, S., 1996. Analysis of the green fraction of humic acids. European Journal of Soil Science 47, 197-204.

Whiteside, G.B., 1950. Soil Survey of Prince Edward Island, in: Agriculture, C.D.o., Agriculture, P.E.I.D.o. (Eds.). Experimental Farms Service.

Woods, G.C., Simpson, M.J., Koerner, P.J., Napoli, A., Simpson, A.J., 2011. HILIC-NMR: Toward the identification of individual molecular components in dissolved organic matter. Environmental Science & Technology 45, 3880-3886.

Woods, G.C., Simpson, M.J., Simpson, A.J., 2012. Oxidized sterols as a significant component of dissolved organic matter: Evidence from 2D HPLC in combination with 2D and 3D NMR spectroscopy. Water Research 46, 3398-3408.

Xu, C., Athon, M., Ho, Y.F., Chang, H.S., Zhang, S.J., Kaplan, D.I., Schwehr, K.A., DiDonato, N., Hatcher, P.G., Santschi, P.H., 2014. Plutonium Immobilization and Remobilization by Soil Mineral and Organic Matter in the Far-Field of the Savannah River Site, US. Environmental Science & Technology 48, 3186-3195.

Xu, C., Santschi, P.H., Zhong, J.Y., Hatcher, P.G., Francis, A.J., Dodge, C.J., Roberts, K.A., Hung, C.C., Honeyman, B.D., 2008. Colloidal Cutin-Like Substances Cross-Linked to Siderophore Decomposition Products Mobilizing Plutonium from Contaminated Soils. Environmental Science & Technology 42, 8211-8217.

Xu, C., Zhang, S., Kaplan, D.I., Ho, Y.-F., Schwehr, K.A., Roberts, K.A., Chen, H., DiDonato, N., Athon, M., Hatcher, P.G., Santschi, P.H., 2015. Evidence for hydroxamate siderophores and other N-containing organic compounds controlling 239,240Pu immobilization and re-mobilization in a wetland sediment. Environmental Science & Technology 49, 11458-11467.

Yuan, T.-Q., Sun, S.-N., Xu, F., Sun, R.-C., 2011. Characterization of lignin structures and lignin–carbohydrate complex (LCC) linkages by quantitative ^{13}C and 2D HSQC NMR spectroscopy. Journal of Agricultural and Food Chemistry 59, 10604-10614.

Zhong, J., Sleighter, R.L., Salmon, E., McKee, G.A., Hatcher, P.G., 2011. Combining advanced NMR techniques with ultrahigh resolution mass spectrometry: A new strategy for molecular scale characterization of macromolecular components of soil and sedimentary organic matter. Organic Geochemistry 42, 903-916.

APPENDIX A

ABBREVIATIONS AND ACRONYMS

Ai_{mod}, *modified aromaticity index;*

BC, blackcarbon;

CCAM, carboxyl containing aliphatic molecules;

CRAM, carboxyl-rich alicyclic molecules;

CPMAS, cross-polarization magic angle spinning;

COSY, correlation spectroscopy;

DBE, double bond equivalents;

DFO, desferrioxamine

DPMAS, direct polarization magic angle spinning;

DOM, dissolved organic matter;

ESI, electrospray ionization;

ESI-FTICR-MS, electrospray ionization coupled to Fourier transform ion cyclotron resonance mass spectrometry;

HA, humic acid;

HMBC, heteronuclear multiple bond coherence;

HSQC, heteronuclear single quantum coherence:

IEF, iso-electric focusing;

KMD, Kendrick mass defect;

LODOM, Lake Ontario dissolved organic matter;

NOM, natural organic matter;

NMR, nuclear magnetic resonance spectroscopy;

N/C, nitrogen to carbon ratio;

Pu, Plutonium;

RFETS, Rocky Flats Environmental Technology Site;

SOM, soil organic matter;

TOCSY, total correlation spectroscopy;

UDOM, ultrafiltered dissolved organic matter

APPENDIX B

COPYRIGHT PERMISSIONS

Permission for Chapter II, which contains the Geochimica et Cosmochimica Acta article,

published by Elsevier, and accessible by the following link

http://www.sciencedirect.com/science/article/pii/S0016703716000387.

This Agreement between Nicole DiDonato ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4037260058022
License date	Jan 27, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	Geochimica et Cosmochimica Acta
Licensed Content Title	Potential origin and formation for molecular components of humic acids in soils
Licensed Content Author	Nicole DiDonato,Hongmei Chen,Derek Waggoner,Patrick G. Hatcher
Licensed Content Date	1 April 2016
Licensed Content Volume Number	178
Licensed Content Issue Number	n/a
Licensed Content Pages	13
Start Page	210
End Page	222
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	ALICYCLIC AND AROMATIC CARBOXYLIC ACIDS IN SOIL ORGANIC MATTER: AN INVESTIGATION OF POTENTIAL ORIGIN AND ASSOCIATION WITH

	PLUTONIUM USING ADVANCED ANALYTICAL TECHNIQUES
Expected completion date	May 2017
Estimated size (number of pages)	150
Elsevier VAT number	GB 494 6272 12
Requestor Location	Nicole DiDonato 4402 Elkhorn Ave NORFOLK, VA 23529 United States Attn: Nicole DiDonato
Total	0.00 USD

Permission for Chapter IV, which contains the Environmental Science and Technology article,

published by ACS, and accessible by the following link

http://pubs.acs.org/doi/abs/10.1021/acs.est.6b04955.

Title:

Substructural Components of Organic Colloids from a Pu-Polluted Soil with Implications for Pu Mobilization

Author:

Nicole DiDonato, Chen Xu, Peter H. Santschi, et al

Publication:

Environmental Science & Technology

Publisher:

American Chemical Society

Date:

Apr 1, 2017

If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials.

Already a RightsLink user or want to learn more?

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.

One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

VITA

Nicole DiDonato

Department of Chemistry and Biochemistry, Old Dominion University, Norfolk, VA 23529

Education
May 2017 ....Ph.D. in Chemistry, Old Dominion University, Norfolk, VA
May 2000 ....Bachelor of Environmental Engineering, University of Delaware, Newark, DE

Publications
DiDonato N, Chen H, Waggoner D, Hatcher PG. (2016) Potential origin and formation for molecular components of humic acids in soils. Geochimica et Cosmochimica Acta 178, 210-222

Xu, C., Zhang, S., Kaplan, D. I., Ho, Y.-F., Schwehr, K. A., Roberts, K. A., Chen, H., DiDonato, N., Athon, M., Hatcher, P.G., Santschi, P. H. (2015) Evidence for Hydroxamate Siderophores and Other N-Containing Organic Compounds Controlling 239,240Pu Immobilization and Remobilization in a Wetland Sediment. Environmental Science & Technology 49(19): 11458–11467

Xu, C., M. Athon, Ho, Y.-F., H.-S. Chang, S. Zhang, D. I. Kaplan, K. A. Schwehr, N. DiDonato, P. G. Hatcher and P. H. Santschi (2014). Plutonium Immobilization and Remobilization by Soil Mineral and Organic Matter in the Far-Field of the Savannah River Site, U.S. Environmental Science & Technology 48(6): 3186-3195

Zhong, J., N. DiDonato and P. G. Hatcher (2012). Independent component analysis applied to diffusion-ordered spectroscopy: separating nuclear magnetic resonance spectra of analytes in mixtures. Journal of Chemometrics 26(5): 150-157

Selected Presentations
DiDonato N, and Hatcher P.G. "Multi-dimensional NMR and ESI-FTICRMS Analysis of soil Humic acid: An investigation of Aliphatic Carboxylic acids." Oral Presentation at the International Humic Substances Society 18th International Conference, Kanazawa, Japan, September 11-16, 2016

DiDonato N, and Hatcher P.G. "Alicyclic Carboxylic Acids in Humic Acid as Detected by Multi-dimensional NMR and ESI-FTICRMS." Poster presentation at the Organic Geochemistry Gordon Research Conference: Holderness, New Hampshire, July 24-29, 2016

DiDonato N., Chen, H., Waggoner, D., Hatcher, P.G. "Soil Humic Acids- a potential common source and formation process elucidated by ESI-FTICR-MS." Poster presentation at the ASMS 63rd Annual Conference, St. Louis MO, May 31-June 5, 2015

Employment Experience
2001-2007   Environmental Engineer, Merck & Co., Inc., Rahway, New Jersey

Selected Honors/Awards
2014   Student Travel Fellowship Award DOE TES/SBR Joint Investigator's Meeting
2012   CIBA Scholarship, Old Dominion University
2007   Excellence Fellowship, Rutgers University
2004-2005   Certificate of Excellence Awards, Merck & Co., Inc.