JOURNAL OF MASS SPECTROMETRY
J. Mass Spectrom. 2007; **42**: 559–574
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/jms.1221

JMS

SPECIAL FEATURE: PERSPECTIVE

The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter

Rachel L. Sleighter and Patrick G. Hatcher*

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

Received 2 February 2007; Accepted 27 March 2007

Mass spectrometry has recently played a key role in the understanding of natural organic matter (NOM) by providing molecular-level details about its composition. NOM, a complex assemblage of organic molecules present in natural waters and soils/sediments, has the ability to bind and transport anthropogenic materials. An improved understanding of its composition is crucial in order to understand how pollutants interact with NOM and how NOM cycles through global carbon cycles. In the past, low-resolution (>3000) mass analyzers have offered some insights into the structure of NOM, but emerging ultrahigh resolution (>200 000) techniques such as electrospray ionization (ESI) coupled to Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) have significantly advanced our knowledge of NOM chemistry. Here, a review of the recent literature on the advancements of NOM characterization and the applications of mass spectrometry to this central task is presented. Various methods for the analysis and display of the extremely complex mass spectra, such as the van Krevelen diagram and Kendrick mass defect analysis, are discussed. We also review tandem mass spectrometry techniques employed to gain structural information about NOM components. Finally, we show how ESI-FT-ICR-MS has been applied to examine specific issues that are important to the NOM scientific community, such as NOM reactivity, transport and fate, degradation, and existence of components, which are indicators of NOM origin. In general, ultrahigh resolution provided by FT-ICR-MS is essential for the complete separation of the thousands of peaks present in the complex NOM mixture and will clearly lead to additional future advancements in the areas of aquatic, soil, and analytical chemistry. Copyright © 2007 John Wiley & Sons, Ltd.

KEYWORDS: electrospray ionization; Fourier transform ion cyclotron resonance mass spectrometry; dissolved organic matter; humic substances; molecular-level characterization

INTRODUCTION

Natural organic matter (NOM) is a complex mixture of organic compounds derived mostly from decaying vegetation in the environment. Present in sediments, soils, and water, these substances have defied molecular-level characterization for nearly a century, owing primarily to the fact that they exist mostly as highly functionalized polyelectrolytes and, as such, do not lend themselves to analytical techniques for molecular characterization. Soluble extracts of NOM can be recovered from soils and sediments by well-established alkaline extraction protocols,[1] and these are referred to as humic substances (HS). NOM that is present in natural waters, namely dissolved organic matter (DOM), is operationally defined as the fraction that will pass through a filter of a nominal pore size (usually between 0.1 and 1.0 microns). NOM and DOM are extremely complex mixtures that can be subcategorized into humic and nonhumic substances. Nonhumic substances are those that retain their chemical characteristics and reactivity as individual compounds within the mixture, i.e. simple sugars, fatty acids, carbohydrates, peptides, etc. HS are the uncharacterized portion of NOM that is composed of a mixture of individual compounds that tend to aggregate into a larger molecular weight and act as a polymeric unit. NOM also tends to be more resistant to microbial degradation. HS can be further classified into humic acids (HA), fulvic acids (FA), and humin. HA is base-soluble and acid-insoluble, FA is soluble across the entire pH range, and humin is insoluble across the entire pH range. Isolation of HA and FA is typically achieved with a simple acid/base extraction procedure.[1,2]

*Correspondence to: Patrick G. Hatcher, 4541 Hampton Blvd, Norfolk, VA 23529, USA. E-mail: phatcher@odu.edu

JMS

NOM contains a variety of components with numerous physical and chemical properties. Polar and nonpolar compounds are present, and although they are water soluble, they still maintain a certain hydrophobic functionality. NOM plays many vital roles in natural waters and soils/sediments, from its ability to bind and transport anthropogenic materials to its part in the numerous biogeochemical processes that mediate our environment. An improved understanding of its composition is necessary to further understand how pollutants interact with NOM. Additionally, riverine NOM is a major source of carbon to the oceans and plays a pivotal role in the global carbon cycle. For a more detailed description of NOM and DOM, along with its formation and transport, see Stevenson.[3]

NOM is well known to provide an analytical challenge to individuals seeking molecular-level information. This is mainly due to its extreme complexity, low concentration, and high polarity. Numerous techniques, such as pyrolysis gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), tetramethylammonium hydroxide (TMAH) thermochemolysis or other thermolysis techniques, and nuclear magnetic resonance (NMR), have been employed to characterize NOM from a variety of sources, but most of these methods are biased and/or fail to fully resolve the numerous constituents in DOM mixtures.[4,5] Low-resolution GC- and LC-MS are not particularly suited for the analysis of NOM, simply because volatility and solubility are required, respectively, allowing these techniques to examine only a small portion of the overall NOM. To employ these methods effectively, NOM must be chemically degraded to smaller molecules that lend themselves to either derivatization and/or volatization. The major disadvantage of pyrolysis and thermochemlysis methods to accomplish this is that both are invasive techniques that can form products that were not initially present in the NOM. Furthermore, pyrolysis and thermochemlysis are very selective in the types of molecules they are capable of analyzing. NMR yields information about functional groups present in the bulk NOM, but fails to give molecular-level details.

Some recent advances have emerged in the characterization of polar, macromolecular molecules dissolved in aqueous solvents. These have come about from the biomedical and biochemical research communities, faced with the daunting task of analyzing proteins, peptides, and metabolic products from organisms, all of which behave like NOM, in that they are generally present in aqueous systems. Thus, molecular characterization methods aimed at obtaining structural information on the myriad of complex biological molecules have evolved and these have been shown to be ideally suited for characterizing NOM molecules. Foremost in these ventures has been mass spectrometry, where determining molecular weights and structural information is possible. The key to the application of mass spectrometry to biological systems was the development of soft ionization methods designed to transform polar analytes to charged species that could be manipulated within the gaseous or high-vacuum phase of mass spectrometers.

Here we will explore the various ionization sources and mass spectrometers that have been utilized for environmental samples, and emphasize particularly how electrospray ionization (ESI) coupled to ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) has succeeded in the extensive characterization of NOM where other methods have been marginally or less successful. In addition, recent applications and their impact on the understanding of NOM compositions are discussed. The purpose of this review is not to describe the instrumentation in detail, but rather to discuss how they are applied to samples of environmental concern, NOM, and utilized to characterize the types of compounds present, at the molecular level.

ELECTROSPRAY IONIZATION

While there are numerous ionization techniques available for mass spectral analysis, ESI is emerging as the optimal method for NOM. This is because ESI takes place at atmospheric pressure, ionizes a wide range of polar, hydrophilic molecules with both acidic and basic functional groups (the kinds of molecules that are abundant in NOM), and can be operated in the positive or negative ion mode. Also, since the samples for ESI are made up in a mixture of water and/or simple, low-molecular-weight organic solvents that evaporate during ESI, solvent-generated ions do not interfere with the mass spectral information generated for the NOM. The mechanism of ESI has been explained in great detail elsewhere,[6] but briefly, a high voltage difference between the needle and capillary causes the sample to burst into many smaller charged droplets. This voltage difference will determine whether positive or negative ions are formed, i.e. the sign of the voltage difference is the opposite of the sign of the ions formed. As the solvent evaporates with the aid of heat or nitrogen gas, the charges on the droplet accumulate and begin to repel each other owing to charge–charge interactions. Once the droplet reaches the Rayleigh limit, where the Coulombic repulsions outweigh the solvent surface tension, the droplet will once again explode into many smaller droplets. The remaining solvent completely evaporates, leaving the analyte molecule in the gas phase, with a charge distributed among the polar functional groups present in the analyte.

Although some researchers have found that ESI can produce multiply charged species for NOM,[7,8] recent investigations have proven that the overwhelming majority of ions produced from the ESI of NOM are singly charged.[9–11] Furthermore, ESI is a 'soft' ionization technique. Accordingly, fragmentation of analyte molecules is basically nonexistent, and only deprotonated molecular ions or molecular ion adducts (H^+ or Na^+) are observed in the mass spectrum. Numerous experiments have been performed to test whether NOM molecules can fragment or dissociate in the ESI source, such that weak bonds or weak interactions like hydrogen bonding and van der Waal forces are broken and macromolecules are split into smaller units.[11] In these experiments, simple polymers can be shown to dissociate readily under the variety of conditions used, but NOM seems to either be

DOI: 10.1002/jms

fully dissociated to the maximum (e.g. all weak noncovalent interactions such as metal ligation or hydrogen-bonding interactions are broken) or little dissociation is observed. To date, there have been few attempts to examine the impact of ESI on structural integrity of NOM molecules other than the work of Rostad and Leenheer[12] and Stenson *et al*.,[11] who examined the relationship of various ESI operating conditions on the spectral characteristics of NOM. Unfortunately, this issue regarding the extent of molecular disruption in ESI spectra of NOM has not been rigorously addressed, in large part because one does not know the exact molecular makeup of NOM to be able to ascertain the extent of disaggregation.

Many investigators have tried other ionization methods, such as electron ionization (EI), pyrolysis, and matrix-assisted laser desorption ionization (MALDI), but these techniques all have their certain biases and disadvantages. For further reading on applications of these ionization methods to NOM and DOM, readers are referred to the several publications.[13–19] Also, for more mechanistic descriptions of any of these ionization methods, see Hoffmann and Stroobant.[20]

Positive ion mode *vs* negative ion mode

As mentioned before, ESI can be operated in both positive and negative ion modes. Depending on the sample, each ionization mode can give very different mass spectra for the same sample,[7,12] and one must consider the complexity that evolves from each ionization mode. Because most investigators are concerned with the ability of a particular mass spectrum to 'represent' the sample being examined, it is important to determine if the ionization process is overly selective for certain types of molecules. If selectivity for one type of molecule is high, then it can be expected that the mass spectrum will not represent all molecules equally. Of paramount importance are ionization efficiencies, which are determined by the types of functional groups present in the sample and their ability to lose or accept a proton (or other cation) for negative or positive ionization, respectively. For example, a sample with easily ionizable and/or numerous acidic groups, such as carboxylic acids, will readily lose a proton and be negatively ionized very efficiently.[21] In contrast, samples with many basic groups, such as amines, will easily pick up a proton and be positively ionized. Since it is known that NOM and DOM samples contain both these types of functionalities, many investigators have used both modes concurrently to characterize samples to the best of their instrument's ability.[7,12] Rostad and Leenheer[12] found that the positive ion mass spectrum of Suwannee River fulvic acid was less effective but more complicated than the negative ion spectrum of the same sample. The positive ion mass spectrum was more complicated because of the introduction of sodium into the sample, giving numerous sodium adducts. When hydrogen and sodium adducts form for the same molecule, for all molecules in the sample, one will observe twice the number of peaks than normally detected if sodium was absent. It should be noted that chlorine adducts can form in the negative ion mode if NaCl is present in the sample because of an incomplete desalting procedure. Also shown by Rostad and Leenheer[12] was that very different mass spectra can be acquired for the same sample from positive and negative ionization modes, indicating that numerous functional groups are present and they are preferentially ionized on the basis of their relative acidity and basicity. Both spectra can be used to give complementary information about the sample, as long as relative abundances are not compared, because of the different ionization efficiencies of the various molecules.

Spray composition considerations

Sample makeup, i.e. concentration, pH, aqueous and organic solvent percentages, and ionic strength, can greatly influence the quality of the mass spectra obtained. Not only do these parameters affect the abundance of ions created in the source region, but also the number of ions detected by the mass spectrometer. In studies of organic modifier concentrations, Kujawinski *et al*.[22] noted that a 25% aqueous solution of a Mount Rainier humic acid extract did not differ from a 50% aqueous solution, when mixed with methanol. In addition, Rostad and Leenheer[12] found that methanol : water mixtures of Suwannee River NOM gave better results than acetonitrile : water and that high-surface-tension solvents such as purely aqueous or dimethyl sulfoxide (DMSO) give unstable spray conditions.

High salt concentrations can easily interfere with the ionization process, because of an unstable spray ('spitting' caused by corona discharge) or salt precipitating in the spray needle, clogging the line. Stenson *et al*.[11] determined that millimolar concentrations of NaOH added to Suwannee River NOM significantly lowered the ion current and signal-to-noise ratios; other investigators found very similar results.[7,22] It is very common to add acid or base (for positive or negative ionization, respectively) to samples prior to ESI in order to increase the number of ions formed in the source region. Brown and Rice[7] performed numerous experiments and concluded that for positive ionization of peat NOM, pH 2 gave the highest quality mass spectra when compared to a pH of 4, 6, or 8. Those higher pH values gave spectra indicative of salt–solvent adduct ions, similar to what was described above with higher ionic strength NOM. Correspondingly, they also found that a higher pH value gave the optimal results for negative ionization.

ATMOSPHERIC PRESSURE PHOTOIONIZATION (APPI)

Although ESI is by far the most common ionization method for NOM and DOM samples, APPI is emerging as a new and improved approach. This technique, explained in detail in other reviews,[23,24] is a soft ionization technique that extends the mass spectrometric analysis of liquid samples to analytes that are not amenable to electrospray ionization. Direct ionization of the analyte (AB) is initiated by electronic excitation (*), which occurs upon the absorption of UV photons ($h\nu$).

$$AB + h\nu \longrightarrow AB^*$$

When $h\nu$ is greater than the ionization energy (IE) of the analyte, the following ionization reaction occurs:

$$AB^* \longrightarrow AB^{+\cdot} + e^-$$

DOI: 10.1002/jms

JMS

A large portion of the organic molecules typically under investigation has IEs less than 10 eV. When using lamps with suitable photon energies (e.g. krypton), the abundant solvent can deplete the photons intended for ionization of analyte molecules. This problem can be overcome by use of a dopant. Because they generate photoions that undergo charge exchange and proton transfer reactions with the analyte, dopants such as toluene and tetrahydrofuran function effectively as intermediates between photons and analytes. In contrast with ESI, nonpolar molecules are also ionized and, thus, detectable by MS. Overall, APPI has a promising future, due to its ability to ionize what was previously thought to be the invisible molecular fraction of NOM, mainly because ESI is not effective in ionizing such molecules (compounds such as nonpolar macromolecules, hydrocarbons, etc.). However, there are no published works, to our knowledge at this time, on the use of APPI for analysis of NOM. Our focus in this review will, accordingly, be on ESI methods that have dominated the literature in the area of NOM studies by mass spectrometry.

LOW-RESOLUTION MASS SPECTROMETRY

Before the development and optimization of high-resolution mass spectrometers such as FT-ICR-MS, NOM and DOM were being examined with lower-resolution instruments; for example, the single quadrupole (Q),[8,12,25–27] the triple quadrupole (QqQ),[28] the quadrupole-time of flight (QqTOF),[10,29,30] the quadrupole ion trap (QIT),[8] and the magnetic/electronic sector (BE).[16,31–33] These mass analyzers operate and acquire data in a multitude of different ways, and this information can be found in any standard mass spectrometry textbook,[20] with more detailed information being found in the literature. All these mass spectrometers give a resolving power of up to approximately 10^4. Unfortunately, as discussed below, this is not nearly the resolution required to resolve all the components in the complex mixtures of NOM and DOM.

Kramer *et al*.[29] utilized QqTOF mass spectrometry to compare HA from an Armadale soil in Canada and a diluvial soil in Japan. They used mass defect, which is the distance a peak is displaced from the exact nominal mass, in the spectra to identify the types of compounds present. Compounds with a low mass defect (<0.1) have low amounts of hydrogen and/or high amounts of oxygen, such as benzocarboxylic acids or carboxylated, condensed aromatics. As the mass defect increases, the number of hydrogens increases and/or the number of oxygens decreases. Mass defects in the range of 0.2–0.4 reflect hydrogen-rich components, such as aliphatic, mid-length fatty acids (C_{15}–C_{22}). Longer-chain fatty acids (C_{24}–C_{40}) appear at mass defects of 0.4–0.6. Even though molecular matches could not be assigned to the peaks in their data set, they found that NOM from various sources have a different makeup at the molecular level based on the mass defects.[29]

Leenheer *et al*.[25] fractionated and analyzed Great Salt Lake DOM using quadrupole mass spectrometry, in combination with other spectroscopic techniques, to fully characterize the DOM and determine its major sources. The polar portion of the sample contained open-chain *N*-acetyl hydroxycarboxylic acids, most likely from polysaccharides in the colloidal material, while the less polar fraction was made up of aliphatic alicyclic ring structures with various functional group substitutions (hydroxyl, carboxyl, methyl, etc.). These findings confirm that the Great Salt Lake DOM is most likely derived from algal and bacterial inputs.[25] Seitzinger *et al*.[26] also utilized quadrupole mass spectrometry to characterize the DOM present in 11 different urban rainwater samples from New Brunswick, NJ; they determined that while there were many compounds unique to each sample, there were also numerous organic acids and bases occurring in the majority of the samples. Quadrupole mass spectrometry was also used to distinguish between DOM from two different streams in New Brunswick, NJ, before and after microbial degradation.[27] Apparently, 70% of the masses detected occurred in both streams and roughly 40–50% of the DOM was bioavailable to the microbes. Furthermore, Seitzinger *et al*.[27] found that the high-molecular-weight fraction was consumed more by the microbes than the low-molecular-weight fraction, and that there was a great deal of agreement between the compounds biologically used in both streams, leading to the belief that microbes' selection of organic molecules for use is consistent within various habitats. Kim *et al*.[34] made similar observations using ultrahigh resolution mass spectrometry where elemental compositions of the molecules could be determined.

Another technique used to detect molecular-level characteristics of DOM is direct temperature-mass spectrometry (DT-MS), which utilizes a magnetic/electronic sector (BE) mass spectrometer. Heat is used to desorb and then pyrolyze DOM compounds, which gives 'fingerprints' of specific compound classes.[16,31] Particulate organic matter (POM) was distinguished from DOM, from both the Delaware Bay and River, by DT-MS, and these studies showed that POM was enriched in proteins, nucleic acids, fatty acids, chlorophyll, and sterols, while DOM was enriched in aminosugars, furfural, and alkylphenol moieties.[32] To further characterize DOM, Minor *et al*.[33] observed that the higher-molecular-weight fraction of ultrafiltered DOM from the lower Chesapeake Bay was rich in aminosugars, deoxysugars, and methylated sugars, whereas the lower-molecular-weight portion was rich in hexose sugars.

To continue this evaluation of mass analyzers, Kujawinski *et al*.[10] compared the QqTOF with the FT-ICR on numerous NOM samples. For each of the different samples, they discovered that the QqTOF was useful in obtaining qualitative differences between samples and the FT-ICR was capable of acquiring molecular-level information. The data acquired on these instruments agree, but their resolution differs quite significantly. Figure 1 shows the differences between the ESI spectra of a single nominal mass region of Suwannee River NOM taken on a QqTOF and one from an FT-ICR-MS. Over most of the entire spectral range (data not shown), the spectra agree quite well, but it is clear that the QqTOF instrument is unable to resolve the complex envelope of peaks observed at each nominal mass region. The advantages and abilities of FT-ICR-MS are discussed further in the next section.

DOI: 10.1002/jms

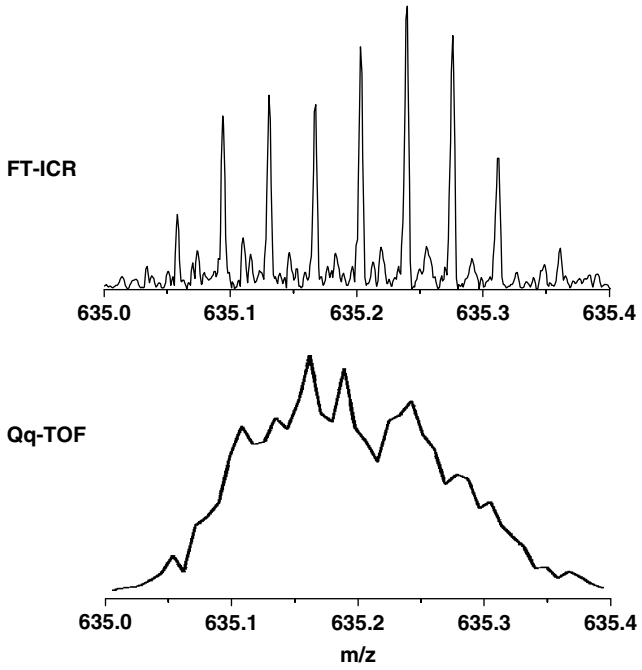


Figure 1. A zoomed-in mass spectrum from 635.0–635.4 to emphasize the differences between the ESI spectra of Suwannee River HS taken on a QqTOF and an FT-ICR-MS. The spectra agree quite well over most of the entire spectral range (data not shown), and it is clear that the instruments correspond in their measurements, but differ in the resolution.

HIGH-RESOLUTION MASS SPECTROMETRY

Although the low-resolution instruments described above can differentiate between trends in mass spectral patterns over a large m/z range or perhaps barely so over one nominal mass range for various NOM samples, the low resolution does not allow the separation of peaks to the level that would be required for differentiation among molecules that vary in mass by less than one mass unit. It is clear from Fig. 1 and many of the articles published recently by our group[34–38] and others,[11,39–42] that NOM is composed of an extremely complex suite of molecules with spectra showing more than 20 peaks per nominal mass over a range of 300 mass units, and only FT-ICR-MS has the capability to resolve the thousands of individual components in the NOM samples. Many investigators have reported resolving powers of in the range of 300 000–600 000.[34,35,39,42]

Marshall *et al*.[43] and references therein, offer an extensive and detailed review of FT-ICR-MS theory and instrumental parameters, but the technique will briefly be discussed here. Basically, FT-ICR-MS is a technique where the ions created in the ESI source are focused and introduced into the horizontal bore of a large cryogenic magnet where they are trapped within a cell. In the presence of a magnetic field, the ions circulate at a frequency that is inversely proportional to their m/z. Positive ions orbit in one direction, while negative ions orbit in the opposite. Frequencies are detected by increasing the diameter of the ion's orbit within the cell using an r.f. pulse; the energy from this pulse is absorbed, which increases the ion's kinetic energy and in turn increases the orbit of the ion within the cell. The packet of ions that is now traveling closer to the detector plates induces an image current on the receiving electrode, and the signal is detected and amplified to give a time domain that is Fourier-transformed into a frequency scale. The frequency scale can be easily converted to the m/z scale by the following simple equation:

$$f_c = zB_o/m$$

where f_c is the cyclotron frequency, z is the charge of the ion, B_o is the strength of the magnetic field, and m is the mass of the ion. Because these frequencies can be measured very accurately, m/z can be determined with very high resolution and precision, usually to the fifth decimal place. With careful calibration, accurate m/z values can be calculated for each peak, allowing the determination of elemental formulas that can be assigned to within 1 ppm error.

Meticulous mass calibration is required to obtain exact and highly resolved m/z values that can be matched with molecular formulas within a 1 ppm error limit or less. Typical external calibration is performed with calibrants such as arginine clusters or a manufacturers' specific tuning mix and can usually be accomplished with an accuracy of 5 ppm. In order to achieve 1 ppm accuracies, those required to obtain reliable elemental formula assignments, internal calibration must be carefully executed. Many studies involving the instrumentation at the National High Magnetic Field Laboratory in Tallahassee, FL, employ a dual-spray injection technique[44] to coinject standards into the source.[34–36] The ions from the standards are accumulated with analyte ions in the hexapole, and both standard and analyte ions are then transferred to the FT-ICR cell. The peaks in the resulting mass spectrum are calibrated by reference to the exact m/z of the internal calibrant ions. Next, only the sample is analyzed by FT-ICR-MS, and the spectra obtained in this step are free from peak contributions of the calibrant material. This spectrum is then internally calibrated by the exact m/z values of the major sample peaks in the spectra obtained in the first step. This approach allows accurate internal calibration, but internal calibration to less than 1 ppm is also possible without the dual-spray procedure. Basically, the same protocol is used, but the sample is mixed with the internal standard prior to electrospray and then analyzed separately without the standard. Also, internal calibration can be achieved with compounds known to be present within the sample. For example, fatty acids are common components of DOM, and because they charge very easily in negative ionization mode,[21] they are ideal for use as internal calibrants. Furthermore, fatty acids are hydrogen-rich and tend to be well separated from other ions typically detected in DOM spectra, making them easily observed.

When FT-ICR-MS was first employed for studies of environmental concern, magnetic fields of 6–7 T were commonplace, offering resolving powers of nearly 80 000.[10,22,45] In these early studies, molecular formulas were not always assigned, because resolving powers did not allow sufficient separation of peaks whose exact masses could be determined with high accuracy. Kim *et al*.[45] used a 7.0 T FT-ICR for a high-DOM stream in the Pinelands of New Jersey and a low-DOM stream in a mountainous region of Costa Rica, and they could

DOI: 10.1002/jms

JMS

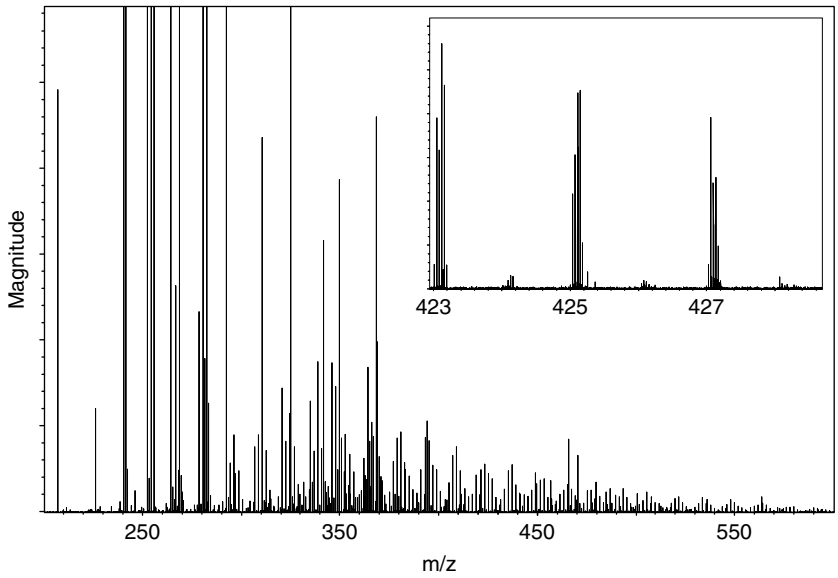


Figure 2. Negative ion 12T FT-ICR mass spectrum of Dismal Swamp DOM, a high-DOC (100 mg/l) swamp water in southeast Virginia and northeast North Carolina. The inset shows the expanded view of *m/z* 423.0–428.5. The isotopic pattern proves that the peaks are singly charged. Formula assignments for the peaks at odd masses are shown in Table 1.

easily distinguish between peaks that were separated by the exact masses of H_2, O, and CH_2. From these observations, they performed Kendrick mass defect analyses[42] (this technique is discussed in detail later) to further examine these different series. Kujawinski *et al.*[22] specifically dedicated a study to determining how resolution can be further increased by FT-ICR mass spectrometers by optimizing certain instrumental parameters. When lower ion densities are injected into the ICR cell, ion–ion interactions (space charge effects) are reduced and signal resolution is enhanced. In addition, coadding larger numbers of transients significantly increases signal-to noise-ratios. These initial studies paved the way for additional research utilizing higher magnetic fields in order to obtain more detailed molecular information, described in a later section.

PRESENTATION OF ULTRAHIGH RESOLUTION MASS SPECTRAL DATA

Molecular formula matches and the van Krevelen diagram

Because of the extremely complex nature of NOM and DOM, analysis by ESI-FT-ICR-MS produces large data sets with thousands of peaks, with as many as twenty peaks at each nominal mass.[9] As mentioned previously, molecular formulas can be determined for each individual peak because the mass can be measured with an accuracy of five decimal places and nearly all peaks are singly charged. Figure 2 shows the negative ion FT-ICR mass spectrum of DOM from the Dismal Swamp obtained at a field strength of 12 T. The inset demonstrates that the ions are in fact singly charged simply because the peaks at the odd masses have their isotopic counterparts located exactly 1.003 mass units higher, the difference between the exact mass of a ^{13}C and a ^{12}C. If the isotope peak is observed at 0.5015 (or 0.3343) higher, then the ion would be doubly (or triply) charged. The molecular formula matches for the peaks at the odd masses from Fig. 2 are shown in Table 1. DBE in Table 1 refers to the Double Bond Equivalents; the equation is shown below.

Table 1. Molecular formula matches for the peaks at odd masses shown as the inset of Fig. 2.
$DBE = 1/2 \times (2\#C + \#N + \#P - \#H + 2)$

Observed *m/z*	Proposed molecular formula	Formula mass	Error (ppm)	*DBE*
423.02048	$C_{17}H_{12}O_{13}$	423.020514	−0.1	12
423.05690	$C_{18}H_{16}O_{12}$	423.056900	0.0	11
423.09326	$C_{19}H_{20}O_{11}$	423.093285	−0.1	10
423.12963	$C_{20}H_{24}O_{10}$	423.129671	−0.1	9
423.15228	$C_{18}H_{24}O_8N_4$	423.152137	0.3	9
423.16601	$C_{21}H_{28}O_9$	423.166056	−0.1	8
423.20240	$C_{22}H_{32}O_8$	423.202442	−0.1	7
425.03601	$C_{17}H_{14}O_{13}$	425.036164	−0.4	11
425.07243	$C_{18}H_{18}O_{12}$	425.072550	−0.3	10
425.08773	$C_{22}H_{18}O_9$	425.087806	−0.2	14
425.10888	$C_{19}H_{22}O_{11}$	425.108935	−0.1	9
425.14532	$C_{20}H_{26}O_{10}$	425.145321	0.0	8
425.18170	$C_{21}H_{30}O_9$	425.181706	0.0	7
425.25767	$C_{20}H_{42}O_7S_1$	425.257848	−0.4	0
425.36366	$C_{26}H_{50}O_4$	425.363634	0.1	2
427.01536	$C_{16}H_{12}O_{14}$	427.015429	−0.2	11
427.05176	$C_{17}H_{16}O_{13}$	427.051814	−0.1	10
427.06697	$C_{21}H_{16}O_{10}$	427.067070	−0.2	14
427.08826	$C_{18}H_{20}O_{12}$	427.088200	0.1	9
427.10352	$C_{22}H_{20}O_9$	427.103456	0.2	13
427.12457	$C_{19}H_{24}O_{11}$	427.124585	0.0	8
427.13987	$C_{23}H_{24}O_8$	427.139841	0.1	12
427.16089	$C_{20}H_{28}O_{10}$	427.160971	−0.2	7
427.19713	$C_{21}H_{32}O_9$	427.197356	−0.5	6

DOI: 10.1002/jms

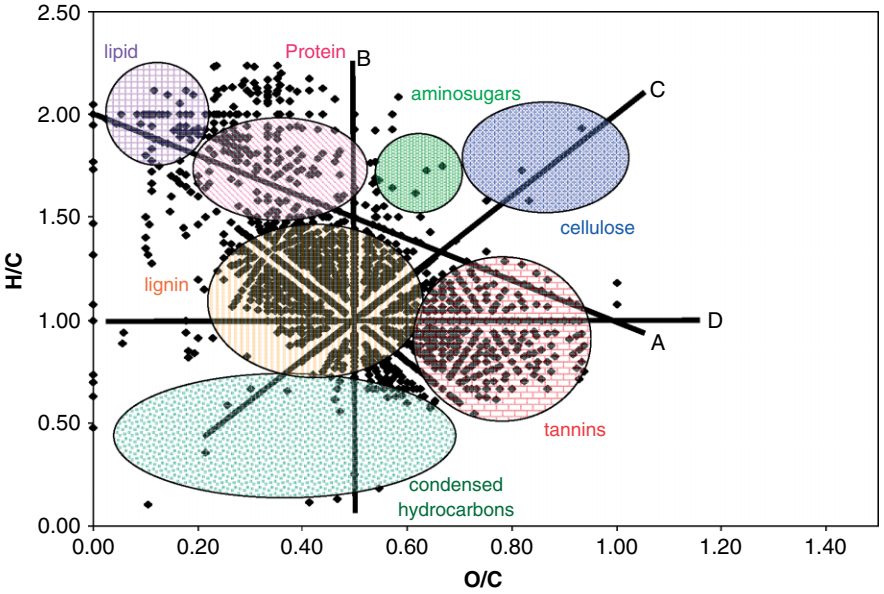


Figure 3. The van Krevelen diagram for the Dismal Swamp DOM, from the molecular formulas calculated from the data shown in Fig. 2. Compound classes are represented by the circles overlain on the plot. The distinctive lines in the plot denote the following chemical reactions: (A) methylation/demethylation, or alkyl chain elongation; (B) hydrogenation/dehydrogenation; (C) hydration/condensation; and (D) oxidation/reduction.

$$DBE = 1/2 \times (2\#C + \#N + \#P - \#H + 2)$$

DBE indicates the total number of double bonds and rings in a molecule. In order to discern the types of molecules present in the sample, one can use the molecular formula matches to construct a van Krevelen diagram. The van Krevelen diagram, first introduced by Van Krevelen[46] and first used for FT-MS data by Kim *et al.*[9] plots the molar H/C ratios on the *y*-axis and the molar O/C ratios on the *x*-axis. Such a plot allows one not only to elucidate what compound classes are present but also to identify what reaction pathways are taking place.[9] Figure 3 shows the van Krevelen diagram created from the formula assignments of the Dismal Swamp DOM mass spectrum shown in Fig. 2, with the major classes of compounds found in NOM overlain on the plot. In addition to clustering the molecules according to their compound classes, well-defined trend lines are observed in the data set, and these are representative of homology among the molecules or can even reflect reaction pathways among sets of molecules. Kim *et al.*[9] explained how masses in this complex mixture of DOM can be related by numerous chemical transformations (e.g. methylation, hydrogenation, hydration, redox, carboxylation, etc.). In a study of the DOM from McDonalds Branch, a blackwater stream in the Pine Barrens area of New Jersey, Kim *et al.*[36] used the van Krevelen diagram to identify peaks in the mass spectrum that could be related to compounds one might expect from combustion of fossil fuels or biomass (i.e. black carbon). Hertkorn *et al.*[47] used this type of plot to identify carboxylic-rich alicyclic molecules in marine DOM, and Stenson *et al.*[42] used the plot to argue that degraded lignin comprises the major precursor to riverine DOM. The van Krevelen diagram has also been employed to differentiate various types of molecules in petroleum and coal extracts.[48,49]

In addition to the two-dimensional van Krevelen diagram shown in Fig. 3, one can also display the information in three dimensions by adding ion abundance or another molar ratio (N/C, S/C, etc.) as the *z*-axis (Figs. 4 and 5). Plotting ion abundance as the third dimension provides an indication of which compound class is in highest abundance, but because ionization efficiency plays a large role in determining the ion's abundance, this comparison should only be used qualitatively. Plotting the *z*-axis as either an N/C ratio or S/C ratio disperses the elemental composition information into a third dimension where one can examine the H/C and O/C ratio of N- or S-containing molecules separated from the clusters of molecules containing only C, H, and O. If peptides are abundantly present, then the N/C ratio will highlight their presence, as shown in Fig. 5. This figure shows the three-dimensional van Krevelen diagram, with the N/C ratio on the *z*-axis, for a humic acid extract of an algal sapropel from Mangrove Lake, Bermuda, a sediment that has been shown to be rich in peptides.[50,51] N/C ratios of 0.0–0.1 suggest long-chain alkyl amines, while N/C ratios of 0.1–0.4 suggest peptides and proteins. Peaks that can be confirmed as originating from peptides are shown in pink on the 3D van Krevelen plot. By and large, these two- and three-dimensional van Krevelen diagrams greatly assist in visualizing the complicated mass spectra that are acquired during the analysis of NOM and DOM samples.

Kendrick mass defect analysis and the *z*-series

When assigning molecular formulas, an important component of most studies of NOM to date, Kendrick mass defect analysis (KMD) can be very useful and has been used extensively.[9,22,37,38,42,45,47,48,52,53] Initial formulas are dispensed by a molecular formula calculator, but as the m/z increases, the number of formulas that match an exact m/z value within a given error value also increases. Typically for

DOI: 10.1002/jms

JMS

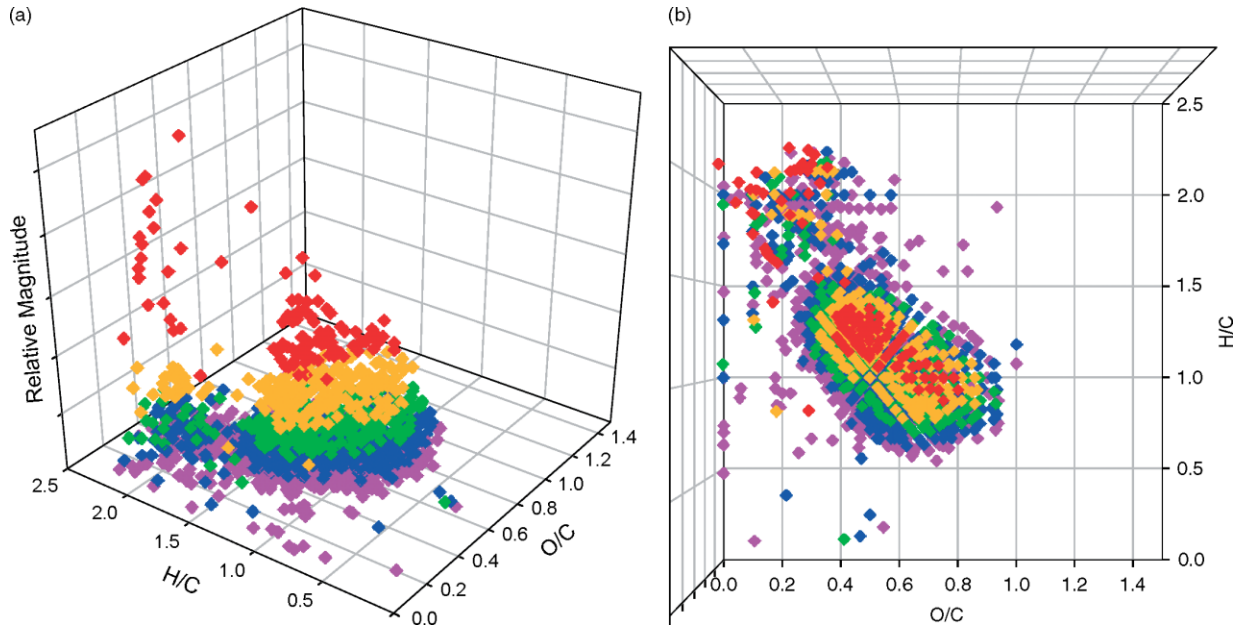


Figure 4. The three-dimensional van Krevelen diagram of Dismal Swamp DOM adding the peak magnitudes as the *z*-axis display (a) and a color-coded plan view (b). Colors of points were varied according to relative peak magnitudes. The magnitudes increase in the order pink, blue, green, orange, and red.

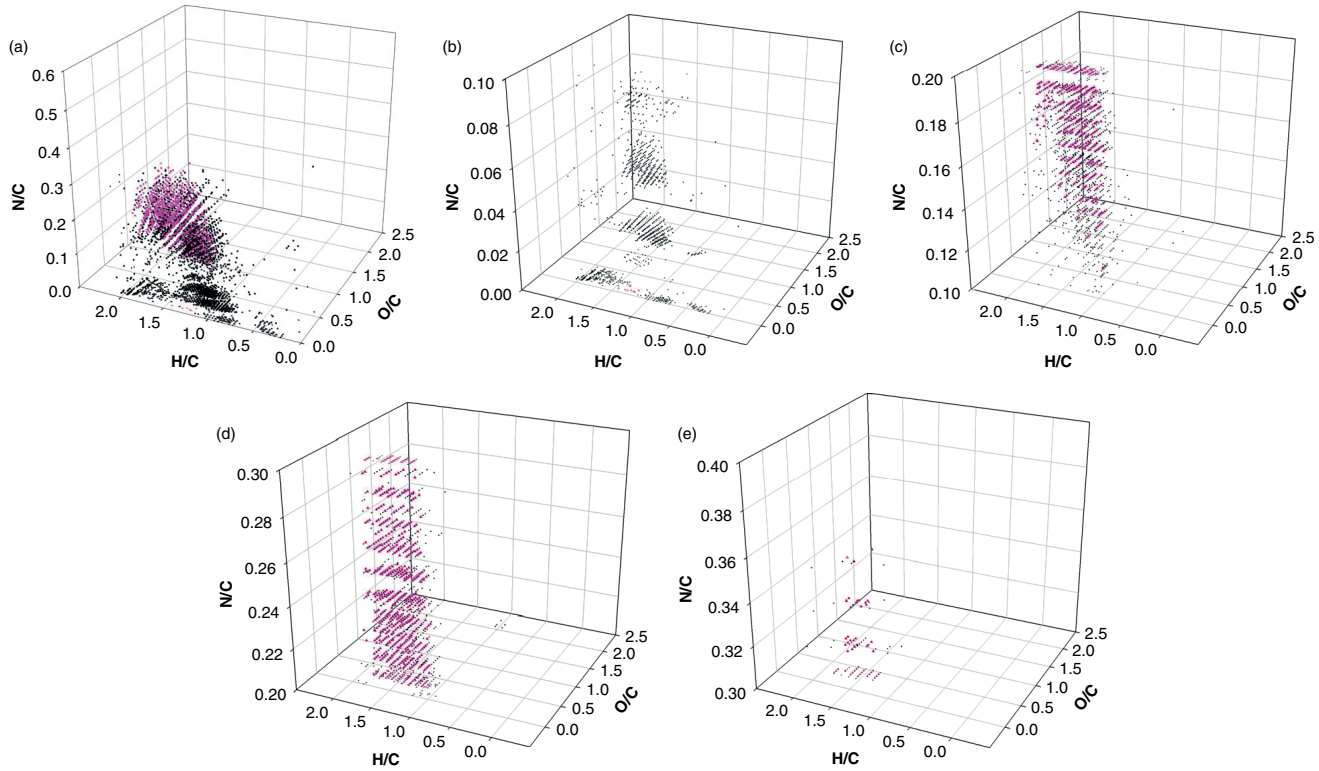


Figure 5. The three-dimensional van Krevelen diagram for a humic acid extract of Mangrove Lake sediment, with the N/C atomic ratio as the *z*-axis (a). Expanded regions are shown for the N/C ratios of 0.0–0.1 (b), 0.1–0.2 (c), 0.2–0.3 (d), and 0.3–0.4 (e). Peaks that can be positively identified as originating from peptides are shown in pink.

low m/z values, a single formula is found within the 1 ppm error window commonly used. However, when more than one formula is extracted within 1 ppm error, the decision needs to be made as to which one is correct. This is where KMD analysis can be utilized, because it identifies homologous series of compounds within a sample. A homologous series is a series of m/z values that differ only by the exact mass of a certain functional group, such as a CH_2 group.[52] KMD basically converts the nominal mass of a CH_2 group (14.000 00) to the exact mass (14.015 65). By multiplying the ratio of the nominal mass to the exact mass (14.000 00/14.015 65) by the m/z value from the mass spectrum, one gets the

DOI: 10.1002/jms

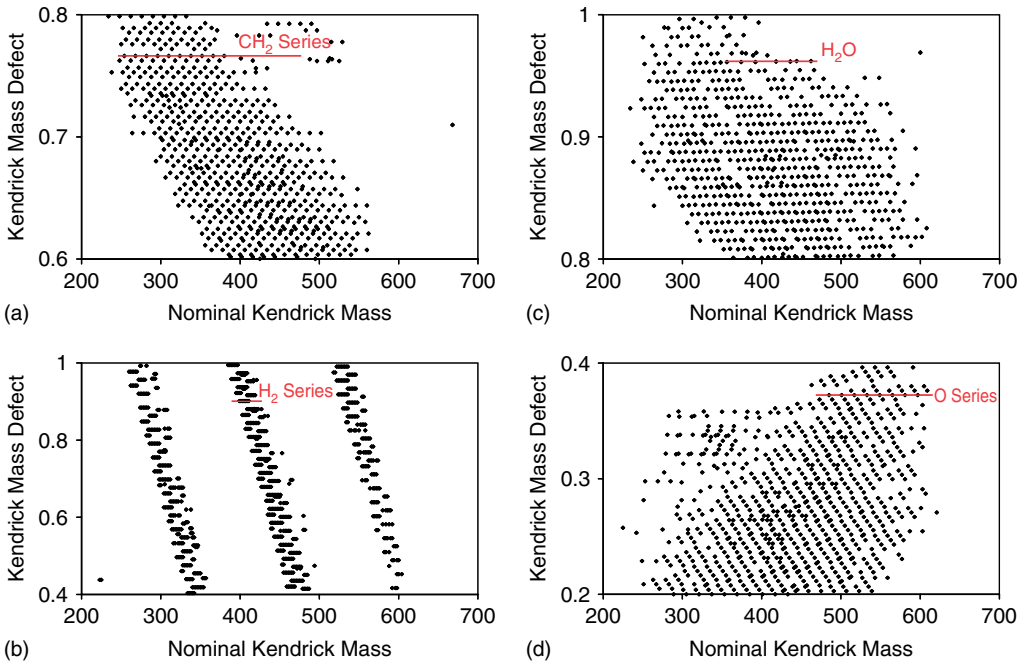


Figure 6. Kendrick mass defect plots of peaks on the trend lines identified in Fig. 3: (a) points in trend line A analyzed by CH_2 KMD analysis; (b) points in line B by H_2 KMD analysis; (c) points in line C by H_2O KMD analysis; and (d) points in line D by O KMD analysis.

Kendrick mass, as shown by the equation below:

$$\text{Kendrick Mass} = \text{exact } m/z \text{ of peak} \times (14.000\,00/14.015\,65)$$

Then, the KMD is calculated by subtracting the Kendrick mass from the observed nominal mass:

$$\text{KMD} = \text{Observed nominal mass} - \text{Kendrick mass}$$

Therefore, ions differing by only a CH_2 group (i.e. chain elongation) will have the same KMD. As a result, if there are multiple formulas for one m/z value, the one that falls within a homologous series is chosen as the correct assignment. By plotting KMD against the nominal Kendrick mass, each homologous series falls on the same horizontal line, as shown by Fig. 6. In this figure, we used KMD analysis to suggest that the chemical reactions from Fig. 3 are occurring. Depending on the makeup of the sample, KMD analysis using other functional groups (e.g. OCH_2, COO, CO, etc.) can be more valuable. Kramer *et al.*[37] used KMD analysis with the COO group to identify black carbon (BC) derived structures in an HA extract of a soil from Japan. With this technique, they were able to distinguish between various types of BC-like structures (discussed further in the BC section). Stenson *et al.*[42] used KMD analysis of Suwannee River FA to assign each m/z value to 1 of 266 unique homologous series, with the assistance of another calculated parameter called the z^* score. z^* is a nominal mass series that helps to differentiate between very close KMD values, so that when ions are assigned to a homologous series, they definitely differ by the exact mass of the functional group that was used.[42] Kujawinski *et al.*[38] also used the z^* approach to separate out changes that occurred to a DOM sample after various levels of degradation had taken place. Thus, not only does KMD analysis assist in assigning molecular formulas to higher m/z ions but it also can aid in the detection of minute changes in the composition of environmental samples after different chemical treatments or various levels of degradation, both aimed at determining specific reaction pathways.

Compacting displays

As already mentioned numerous times and as shown in Fig. 2, FT-ICR-MS produces thousands of peaks in one mass spectrum of an NOM sample. The data acquired from these complex mixtures are not only difficult to visualize but they are also challenging to analyze. Although the van Krevelen diagram and KMD analysis discussed above alleviate most of these problems, the mass spectrum is still difficult to peruse in order to determine the sample's composition. Hughey *et al.*[48] reported that it would be necessary to stretch a mass spectrum out over 200 m in order to see each and every individually resolved peak in that spectrum. Fortunately, Hughey *et al.*[48] designed an extremely beneficial, compact display of a mass spectrum that allows the three-dimensional presentation of all the peaks at each nominal mass. The x-axis is the observed nominal mass, the y-axis is the observed mass defect, and the z-axis is the ion signal magnitude. In this display, one can observe peaks clustered at specific mass defects, and these clusters provide clues as to the types of structurally similar molecules present in the sample. For example, peaks clustered at low mass defects suggest that the sample consists mainly of condensed aromatics, while peaks clustered at a higher mass defect suggest more contributions from hydrogen-rich lignin or lipids. Hence, not only does this compact display assist in viewing all the peaks in a mass spectrum but it can also provide the investigator preliminary results as to the general sample makeup.

DOI: 10.1002/jms

JMS

Difference mass spectra

When comparing two very similar samples, which perhaps have been prepared differently to highlight changes in chemical composition, it is often useful to calculate a difference spectrum to visualize the variations. A reference spectrum can be used as a control and the analyte spectrum can be subtracted from it. Rodgers *et al*.[54] provide details on the computer program used for such a calculation, but we briefly describe it here. The program searches for a peak in the control spectrum and calculates its relative abundance. It then searches for that same peak in the analyte spectrum, using a user-defined mass window, and subtracts its abundance from that of the control spectrum. Positive difference peaks indicate that the assigned constituents are more abundant in the analyte spectrum than in the control spectrum, whereas negative peaks reflect the opposite. Rodgers *et al*.[54,55] used this method to compare a weathered jet fuel, an unweathered jet fuel, and a jet-fuel-contaminated soil. This technique allowed not only increased evaluation of the effect of weathering on a jet fuel but also identification of jet fuel contaminants in the soil. This method can be applied to numerous studies in which it is possible to compare two or more samples.

Automated approaches

Stenson *et al*.[42] reported that it required several months to assign nearly 5000 molecular formulas to the m/z values produced for Suwannee River FA. However, once a database is created, the time to assign molecular formulas for other samples can be significantly reduced. Kujawinski and Behn[53] recognized the need for an automated system so that a high sample throughput could be achieved. They created an automated compound identification algorithm (CIA) that loads and sorts MS data, searches for relationships between the m/z values, and finally assigns molecular formulas to the m/z values on the basis of the functional group relationships established in the previous step. This algorithm was the first attempt at this type of computerized sample analysis. The main obstacle appears to occur with m/z values greater than 500, owing to the larger number of formulas that could be assigned within a specified error from the exact formula mass. If the m/z value could not be correlated to a smaller molecular weight peak by KMD, then CIA is not able to assign a molecular formula.[53] Overall, mass spectral data must be of the highest quality in order for CIA to work efficiently, and careful calibration must be performed in order to have ~95% of m/z values matched with a molecular formula.[53] Their design of an automated CIA has made a significant contribution to the scientific community that is interested in analyzing numerous NOM samples in a reasonable amount of time.

TANDEM MASS SPECTROMETRY

Tandem mass spectrometry (MS/MS) is employed to obtain structural information on specific masses in a spectrum. Typically, the first mass spectrometer selects an ion of specific mass and the second analyzes fragments of that ion induced by a multitude of methods, either bombardment with neutral molecules, electrons, or photons. Thus, one obtains MS/MS data that can be used to infer structure. While this approach works well for mixtures of compounds that are not complex, it has a limited application for NOM studies to date because the first mass spectrometer does not have sufficient mass resolution to isolate an ion from the complex mixture of peaks observed at each nominal mass.

Instrumentation

As mentioned in a previous section, MS/MS can be performed on a QqQ, a QqTOF, a QIT, and an FT-ICR mass spectrometer, and is used to obtain structural information via molecule fragmentation. QqQ and QqTOF are examples of MS/MS being performed by the coupling of two separate instruments. The QqQ and QqTOF mass analyzers combine quadrupole mass filters with either a third quadrupole or a TOF mass analyzer, the latter being aligned orthogonally. The orthogonal geometry of the TOF is essential because it allows the combination of two different types of analyzers: the continuous quadrupole and the pulse-type TOF analyzer. In the quadrupole, ions are continuously traveling through the quadrupole, while TOF analyzers use a pulsed accelerating voltage to allow packets of ions to go through the flight tube, one packet at a time. In both QqQ and QqTOF spectrometers, the first quadrupole selects a certain mass, the precursor ion, that is of interest, and eliminates all other m/z values. The second quadrupole is a collision cell that operates in r.f.-only mode, where an inert or reactive gas is introduced to induce fragmentation. This technique is called collision induced dissociation (CID), and is discussed further in the next section. The TOF (for the QqTOF) or third quadrupole (QqQ) then determines which ions are produced, referred to as product ions, because of the fragmentation. The QIT and FT-ICR mass spectrometers execute MS/MS experiments by performing a sequence of events within the ion storage trap. Since ions can survive several series of excitation events, multiple tandem mass spectrometry experiments can be performed efficiently in the trap. Once an ion is selected and all other ions are ejected from the trap, helium is introduced as a collision gas, since it is already present as a trap coolant. For FT-ICR-MS, an a.c.voltage can be applied to discharge ions of a specified m/z value by making their cyclotron orbits so large that they collide with the wall of the trap. Thus the residual ion(s) can be fragmented by either CID or infrared multiple photon dissociation (IR-MPD), which is discussed in the next section. Unlike QIT MS/MS, FT-ICR-MS/MS detects the fragmented ions in a nondestructive manner, making them detectable repeatedly, increasing sensitivity and resolution. The instrumentation used in these aforementioned studies has only been briefly described here; readers are once again referred to Hoffmann and Stroobant[20] for further descriptions.

Collision induced dissociation *vs* infrared multiple photon dissociation

Overall, CID and IR-MPD both follow the same sequence of events: (1) Isolate precursor ion by ejecting all other ions; (2) Excite the precursor ion; (3) Allow the ion to fragment during a certain reaction period; (4) Excite the product ions;

DOI: 10.1002/jms

(5) Detect the product ions. The only difference is the method used for fragmentation. In order for an ion to fragment, energy needs to be imparted to the ion so that it dissociates. As mentioned in previous sections, CID uses either an inert gas or a reactive gas to induce fragmentation. If an inert gas is used, part of the ion's kinetic energy is converted to internal energy, and low energy fragmentation of that ion occurs. Ion–molecule reactions are generated when a reactive gas is used. Although helium can be used, heavier gases like argon or xenon transfer more energy, thus allowing more collisions to occur. Rather than using a gas to add energy to an ion, IR-MPD utilizes a CO_2 laser to irradiate the ion. IR-MPD has several advantages over CID; for example, the energy added to the ion is more easily controlled by regulating the laser power, and ion dissociation efficiency is much higher. Thus there is no need to introduce another substance that can change the pressure of the high-vacuum system. These benefits lead to a greater sensitivity and selectivity for MS/MS experiments, which is critical for obtaining structural information about the precursor ion. More detailed descriptions of these techniques is beyond the scope of this paper; for further reading, readers are referred to the Refs 20, 56 and 57.

Structural information

Because a mass spectrometer cannot distinguish structural isomers of the same mass, a structure cannot usually be assigned to a specific m/z value. By employing MS/MS via CID or IR-MPD, one can examine the fragmentation patterns of a specific ion and determine the type of structure that is present in abundance. Hoffmann and Stroobant[20] give an excellent overview of how to interpret mass spectra to discern structures from fragmentation pathways. Here, we highlight a few examples of how MS/MS has been applied to NOM samples. Leenheer *et al*.[8] performed CID experiments on Suwannee River FA with a QIT, and losses of H_2O, COO, and CO were observed. These neutral mass losses correspond to alcohol dehydration, decarboxylation, and ester fragmentation, respectively. Furthermore, polycarboxylic acids were used as model compounds to confirm that these reactions were actually taking place. Similar results were obtained by Stenson *et al*.[42] (2003) with IR-MPD of Suwannee River FA using an FT-ICR-MS. They determined that the same neutral losses of H_2O, COO, and CO indicate that the FA structure must consist of a mixture of carboxylic, carbonyl, hydroxyl, and aromatic functional groups. Suwannee River FA were examined by Fievre *et al*.[13] by directly comparing IR-MPD with CID on an FT-ICR-MS. They found that IR-MPD gave a higher abundance of ions and that the fragmentation was more easily controlled by lowering the irradiation period or the laser power. Nonetheless, the more extensive fragmentation seen with CID leads to a more complete understanding of the type of structure that makes up the precursor ion. Plancque *et al*.[30] employed a QqTOF to take these analyses a step further and draw a possible structure for aquifer FAs. Utilizing CID, they explored the fragmentation pathways to develop a structure with numerous carboxylic and hydroxyl groups, with a small number of aromatic groups and very few or no carbonyl groups. The lack of carbonyl groups is due to the fact that they did not observe many CO losses in their tandem mass spectra. To further scrutinize this scheme of numerous carboxylic acid functionalities, McIntyre *et al*.[28] analyzed peat and soil FAs by CID with a QqQ mass spectrometer. The losses of successive COO groups led them to test fragmentation patterns with model compounds containing many carboxyl groups along with other functionalities. They determined that benzene, phenol, dihydroxy benzene, furan, and thiophene carboxylic acids are present in the peat and soil FA and can be readily identified with these MS/MS techniques. As shown, these fragmentation studies can be exploited to give structural information about the analytes, and further investigation of a wide range of NOM and DOM samples will lead to a better understanding of the structures existing in these environmental samples.

THE IMPACT OF HIGH-RESOLUTION MASS SPECTROMETRY ON THE KNOWLEDGE OF THE COMPOSITION OF NOM

The development of ultrahigh resolution MS, in which one can obtain spectrally resolved peaks with sufficient mass accuracy to calculate unique elemental formulas, has revolutionized the study of NOM chemistry and is likely to provide future developments that are destined to change the field dramatically. Accordingly, a number of research groups, including our own, have launched numerous studies designed to take advantage of the offerings of ESI-FT-ICR-MS. Below, we describe some of the initial studies as they relate to specific goals aimed at understanding the molecular makeup of NOM and its reactivity.

Characterization of sources of NOM

Because of the molecular-level detail that FT-ICR-MS provides, NOM samples of different origin can be compared in order to determine the likely sources of this material. Kujawinski *et al*.[22] compared Suwannee River DOM (SRDOM) to a HA extract of Mount Rainier degraded wood (MRHA). The main differences between the two are their sources and degradation pathways; MRHA is made up of mostly lignin degradation products governed by brown rot fungi, and the SRDOM sample contains components from a peat swamp drainage.[22] They primarily used KMD analysis for comparison of the mass spectral data, and found that the MRHA had a greater aromatic contribution, which is consistent with a source from lignin. This conclusion, that lignin constitutes the major source for SRDOM, is supported by more recent studies of Kim *et al*.[45] and Stenson *et al*.[42]

Llewelyn *et al*.[41] used FT-ICR-MS to examine changes in the composition of dissolved organic phosphorus (DOP) in the Everglades Nutrient Removal (ENR) treatment wetland. DOP was collected and concentrated from along a water flow gradient in the ENR, and it was determined that many high-molecular-weight organic phosphorus compounds persisted in all the sites. This observation means that these compounds are part of the refractory component of the DOP and specific enzymes fail to hydrolyze them in order for them to become biologically available.[39,41] This has large implications not only for the relative reactivity of organic phosphorus

DOI: 10.1002/jms

JMS

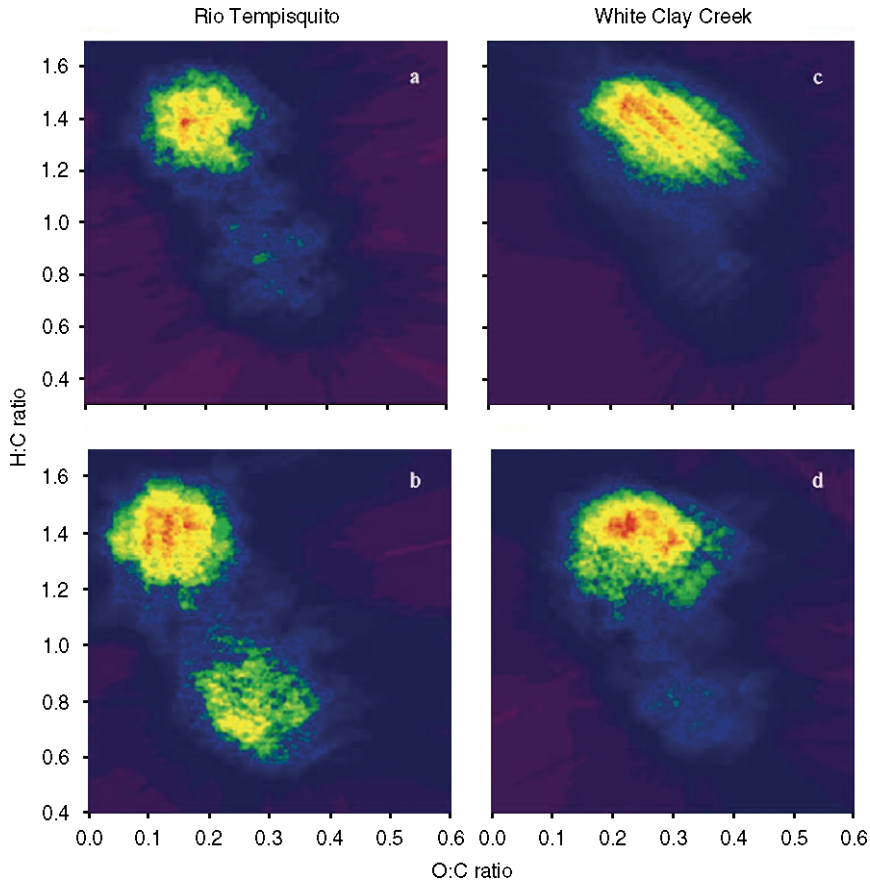


Figure 7. The van Krevelen diagrams of DOM extracted from (a) inflow and (b) outflow of bioreactors at Rio Tempisquito, the low DOC stream and (c) inflow and (d) outflow of bioreactors at White Clay Creek, the higher DOC stream. Reprinted with permission from Kim S, Kaplan LA, Hatcher PG, Biodegradable Dissolved Organic Matter in a Temperate and a Tropical Stream Determined from Ultra-High Resolution Mass Spectrometry, *Limnology and Oceanography*. 2006; 51(2): 1054–1063.[34] Copyright (2006) by the American Society of Limnology and Oceanography, Inc.

compounds but also for insight into phosphorus sources and sinks within the environment. Similar experiments could be designed in order to explore the minute variations of DOM samples along a specific transect, in order to track the transport and eventual fate of particular compounds.

Comparisons of marine *vs* terrestrial DOM

Koch *et al*.[40] examined the differences in a terrestrial DOM sample from mangrove swamp water in Brazil with a marine DOM sample from the Antarctic Sea. Although these water samples are not related, differences in their DOM components could provide a basis for differentiating between oceanic and land-based inputs. Koch *et al*.[40] discovered that, while many (approximately one-third) molecular formulas were identical at both sites, there were many formulas that were unique to each site, with the terrestrial DOM having many more formulas that were different from those observed in oceanic waters. Furthermore, the DOM from various depths in the Antarctic Sea were analyzed and found not to differ significantly from one another. After microbial and photodegradation, the DOM consisted of similar structures that seem to be intrinsically refractory, regardless of source.[40] In another study of marine DOM, Hertkorn *et al*.[47] discovered that carboxyl-rich alicyclic molecules (CRAM) were the most abundantly identified compounds of DOM in the deep ocean (roughly 8% of DOC). This was determined by the complementary information generated by the combined use of FT-ICR-MS and multidimensional NMR. They concluded that polycarboxylated fused ring systems typified CRAM molecules rather than open-chain isomers, and that this CRAM is derived from biomolecules such as sterols and hopanoids.

Elucidation of changes in DOM due to various types of degradation

With ultrahigh resolution FT-ICR-MS, DOM degradation can be discerned if one compares the molecular formulas of peaks before and after degradation. Kujawinski *et al*.[38] exploited this fact by testing the effect of photoirradiation of Suwannee River FA and the function of protozoan grazers in DOM alteration. They noticed that approximately 50% of peaks disappeared after irradiation with long wavelengths. Those peaks missing as the result of degradation are characterized by high DBEs and low oxygen content, and most likely they were modified to a compound that is not easily ionized by electrospray ionization, e.g. a nonpolar hydrocarbon.[38] Furthermore, they observed numerous peaks that were unique to the protozoan grazing incubation that were not present in their bacterial incubation control experiment. This proves that protozoa can make an

DOI: 10.1002/jms

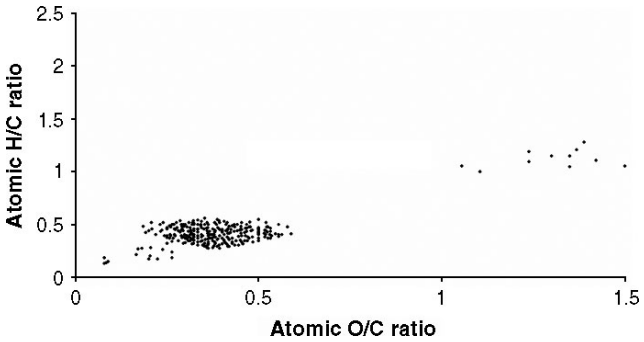
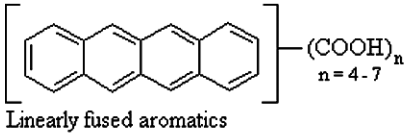


Figure 8. Van Krevelen plot of diluvial humic acid from Japan. Reprinted with permission from Kramer RW, Kujawinski EB, Hatcher PG, Identification of Black Carbon Derived Structures in a Volcanic Ash Soil Humic Acid by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: *Environmental Science and Technology*. 2004; 38: 3387.[37] Copyright (2004) American Chemical Society.

important contribution to DOM by modifying it. In quite a different manner of degradation, Kim *et al*.[34] employed biofilm reactors to determine the transformations of DOM in two different forested stream waters, one with low DOM and the other with a higher DOM concentration. In both streams, biodegradation modifies DOM to lower-molecular-weight constituents, and oxygen-rich molecules were selectively metabolized. In addition, the van Krevelen diagrams shown in Fig. 7 revealed that hydrogen-deficient molecules with low H:C ratios, representative of BC-like structures, are present and generally enhanced in biodegraded waters.[34] They concluded that BC molecules tend to be refractory to biodegradation when compared with other DOM components. This supports the suggestion that BC-derived molecules can flow to the ocean without experiencing significant microbial degradation.

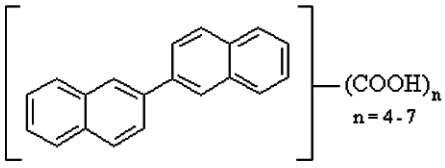
Identification of BC-derived structures

The study of BC's existence in DOM samples has been of particular interest to many investigators, owing to its persistence in the environment. BC is mainly produced by the incomplete combustion of fossil fuels and/or biomass and can be found in many soils, especially those subjected to burning events. Even though BC is thought to be made up of highly condensed aromatic rings, it has recently been thought to somehow become solubilized and incorporated into porewaters, which can then flow into rivers and eventually the ocean.[36,58] Kim *et al*.[36] showed that by use of the van Krevelen diagram, BC peaks could be identified owing to their low mass defect and high DBE values. They speculated that hydroxy groups were substituted onto BC structures, which were then oxidized to quinones or carboxylic acid functionalities. This addition of polar functional groups renders BC more water soluble, explaining its existence in DOM. Kramer *et al*.[37] similarly identified a large suite of BC-derived compounds in an HA extract of a volcanic ash soil from Japan; the van Krevelen diagram strongly indicated the presence of BC, as shown in Fig. 8 They used the KMD analysis, normalized to the COO functional group, in order to identify BC structures varying by the exact mass of COO. By combining the data acquired from the van Krevelen diagram and the KMD analysis, various structural entities were identified, such as linearly fused aromatics, aromatics linked by carbon–carbon single bonds, and highly condensed (fused) aromatic structures. For examples[37] of some of the structures matched with molecular formulas, see Fig. 9. This was the first time that structures could be matched to molecular formulas without the assistance of tandem mass spectrometry. Hockaday *et al*.[35] examined the existence of BC in the soil of a fire-impacted forest along with its pore water, to investigate the solublization of BC and its transport to pore water and adjacent streams. They introduced a new



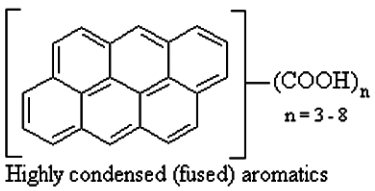
Linearly fused aromatics

Molecular Formula	# of Carboxyls	Error (ppm)
$C_{18}H_{12}$	4	0.97
	5	0.87
	6	0.39
	7	0.09



Condensed aromatic structures linked by biphenyl bonds

Molecular Formula	# of Carboxyls	Error (ppm)
$C_{20}H_{14}$	4	0.63
	5	-0.23
	6	-1.47
	7	-0.50



Highly condensed (fused) aromatics

Molecular Formula	# of Carboxyls	Error (ppm)
$C_{22}H_{12}$	3	0.12
	4	-0.38
	5	0.44
	6	-0.22
	7	0.03
	8	0.00

Figure 9. Linearly fused aromatics, aromatics linked by carbon-carbon single bonds, and highly condensed (fused) aromatic structures. The error represents deviation of the observed mass from the theoretical value in ppm. (Adapted from Kramer *et al*., 2004).

JMS

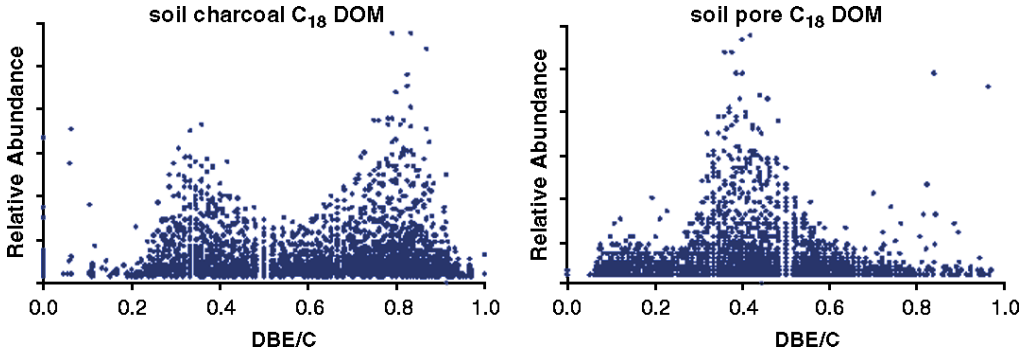


Figure 10. DBE/C spectra of water-soluble charcoal leachates (left) and soil pore water DOM (right) showing that the charcoal is a source of condensed aromatics molecules with DBE/C values distributed around 0.8. Reprinted with permission from Hockaday WC, Grannas AM, Kim S, Hatcher PG. Direct Molecular 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.[35] Copyright (2006) Elsevier.

criterion for identifying condensed aromatic ring structures (CARS), namely, a carbon-normalized DBE value. Thus, BC-derived structures, or CARS, can be differentiated[35] from NOM molecules derived from decaying vegetation if their carbon-normalized DBE exceeds 0.7. A plot of DBE/C *vs* peak abundance provides an estimate of the relative contribution of CARS to the total sample composition, as shown in Fig. 10. Approximately 49% of the peak abundance in a water extract of charcoal residing in soil for about 100 years qualify as CARS, while an average of 10% of the peak abundance in the pore water samples are CARS.[35] This substantiates the fact that BC can be altered to water-soluble compounds and thus exported to the DOM pool.

CONCLUSIONS

ESI coupled to ultrahigh resolution FT-ICR-MS has facilitated the extensive molecular-level characterization of NOM. The advent of ESI has allowed the ionization of the large, nonvolatile compounds, and its application to the polar, polyelectrolytic NOM has made a significant contribution to the understanding of its composition and reactivity. The high resolving power of FT-ICR-MS is capable of separating m/z values to the fifth decimal place, from which molecular formula assignments can be made fairly reliably. Several data analysis and data presentation techniques have been utilized to clearly represent the complicated mass spectra that are acquired during NOM analyses. The most recent applications of ESI-FT-ICR-MS have included DOM characterization, DOM structural features as a function of terrestrial *vs* marine systems, DOM compositional changes as a function of numerous degradation pathways, and the identification of BC in the environment. FT-ICR-MS is clearly a powerful technique used to examine the complex composition of NOM and DOM, and it will most certainly assist in further advancements in the areas of aquatic, soil, and analytical chemistry as individual researchers begin to implement the technique more routinely. The unfortunate aspect of this application is that FT-ICR-MS instrumentation is not readily available to most soil scientists or geochemists, being mainly used for biomedical research applications. However, this situation is slowly changing and we anticipate that access to FT-ICR-MS will become more common.

Acknowledgements

We thank the large number of individuals in the Hatcher research group who contributed indirectly to the work presented here (Dr Sarah A. L. Caccamise, Dr Amanda M. Grannas, Dr Sunghwan Kim, Dr William C. Hockaday, Ms Eunjee Lee, and Mrs Susan A. Hatcher). We thank Dr Alan Marshall and Dr Ryan Rodgers at the National High Magnetic Field laboratory, Florida State University, for making available their 9.4 T FTICR-MS (funded through their NSF Grant CHE-9903528). We thank the support of the National Science Foundation (OCE-0453743).

REFERENCES

1. Schnitzer M, Khan SU. *Developments in Soil Science, Vol. 8: Soil Organic Matter*. Elsevier: Amsterdam, 1978.
2. Thurman EM, Malcolm RL. Preparative isolation of aquatic humic substances. *Environmental Science and Technology* 1981; **15**: 463. DOI: 10.1021/es00086a012.
3. Stevenson FJ. *Humus Chemistry: Genesis, Composition, Reactions*. John Wiley & Sons: New York, 1994.
4. Hatcher PG, Dria KJ, Kim S, Frazier SW. Modern analytical studies of humic substances. *Soil Science* 2001; **166**: 770. DOI: 0038-075C/01/16611-770-794.
5. Leenheer JA, Croué JP. Characterizing aquatic dissolved organic matter. *Environmental Science and Technology* 2003; **37**: 18A.
6. Gaskell SJ. Electrospray: principles and practice. *Journal of Mass Spectrometry* 1997; **32**: 677. DOI: 10.1002/(SICI)1096–9888(199707)32:7<677::AID-JMS536>3.0.CO;2-G.
7. Brown TL, Rice JA. Effect of experimental parameters on the ESI FT-ICR mass spectrum of fulvic acid. *Analytical Chemistry* 2000; **72**: 384. DOI: 10.1021/ac9902087.
8. Leenheer JA, Rostad CE, Gates PM, Furlong ET, Ferrer I. Molecular resolution and fragmentation of fulvic acid by electrospray ionization/multistage tandem mass spectrometry. *Analytical Chemistry* 2001; **73**: 1461. DOI: 10.1021/ac0012593.
9. Kim S, Kramer RW, Hatcher PG. Graphical method of analysis of ultrahigh-resolution broadband mass spectrometry of natural organic matter, the van krevelen diagram. *Analytical Chemistry* 2003; **75**: 5336. DOI: 10.1021/ac034415p.
10. Kujawinski EB, Freitas MA, Zang X, Hatcher PG, Green-Church KB, Jones RB. The application of electrospray ionization mass spectrometry (ESI MS) to the structural characterization of natural organic matter. *Organic Geochemistry* 2002; **33**: 171. DOI: 10.1016/S0146-6380(01)00149-8.
11. Stenson AC, Landing WM, Marshall AG, Cooper WT. Ionization and fragmentation of humic substances in electrospray

ionization fourier transform ion cyclotron resonance mass spectrometry. *Analytical Chemistry* 2002; **74**: 4397. DOI: 10.1021/ac020019f.
12. Rostad CE, Leenheer JA. Factors that affect molecular weight distribution of suwannee river fulvic acid as determined by electrospray ionization/mass spectrometry. *Analytica Chimica Acta* 2004; **523**: 269. DOI: 10.1016/j.aca.2004.06.065.
13. Fievre A, Solouki T, Marshall A, Cooper WT. High resolution fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids using laser desorption and electrospray ionization. *Energy and Fuels* 1997; **11**: 554. DOI: 10.1021/ef97005q.
14. Leinweber P, Schulten HR. Composition, stability and turnover of soil organic matter – investigations by off-line pyrolysis and direct pyrolysis mass spectrometry. *Journal of Analytical and Applied Pyrolysis* 1995; **32**: 91. DOI: 10.1016/0165-2370(94)00832-L.
15. Leinweber P, Schulten HR. Advances in analytical pyrolysis of soil organic matter. *Journal of Analytical and Applied Pyrolysis* 1999; **49**: 359. DOI: 10.1016/S0165-2370(98)00082-5.
16. Minor EC, Eglinton TI, Boon JJ. Direct temperature-resolved mass spectrometry as a technique for the semi-quantitative analysis of marine particulate organic matter. *Journal of Analytical and Applied Pyrolysis* 2000; **53**: 19. DOI: 10.1016/S0165-2370(99)00053-4.
17. Remmler M, Georgi A, Kopinke FD. Evaluation of matrix-assisted laser desorption/ionization (MALDI) time-of-flight (TOF) mass spectrometry as a method for the determination of the molecular mass distributions of humic acids. *European Journal of Mass Spectrometry* 1995; **1**: 403. DOI: 10.1255/ejms.96.
18. Saiz-Jimenez C. Analytical pyrolysis of humic substances: pitfalls, limitations, and possible solutions. *Environmental Science and Technology* 1994; **28**: 1773. DOI: 10.1021/es00060a005.
19. Schulten HR, Leinweber P. Characterization of humic and soil particles by analytical pyrolysis and computer modeling. *Journal of Analytical and Applied Pyrolysis* 1996; **38**: 1. DOI: 10.1016/S0165-2370(96)00954-0.
20. Hoffmann E, Stroobant V. *Mass Spectrometry: Principles and Applications*. John Wiley & Sons: Chichester, 2003.
21. Lee EJ. Studies of dissolved black carbon and organic matter utilizing electrospray ionization mass spectrometry. M.S. thesis. The Ohio State University. 2005.
22. Kujawinski EB, Hatcher PG, Freitas MA. High-resolution fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids: improvements and comparisons. *Analytical Chemistry* 2002b; **74**: 413. DOI: 10.1021/ac0108313.
23. Bos SJ, Van Leeuwen SM, Karst U. From fundamentals to applications: recent developments in atmospheric pressure photoionization mass spectrometry. *Analytical and Bioanalytical Chemistry* 2006; **384**: 85. DOI: 10.1007/s00216-005-0046-1.
24. Raffaelli A, Saba A. Atmospheric pressure photoionization mass spectrometry. *Mass Spectrometry Reviews* 2003; **22**: 318. DOI: 10.1002mas.10060.
25. Leenheer JA, Noyes TI, Rostad CE, Davisson ML. Characterization and origin of polar dissolved organic matter from the great salt lake. *Biogeochemistry* 2004; **69**: 125. DOI: 10.1023/B:BIOG.0000031044.16410.27.
26. Seitzinger SP, Styles RM, Lauck R, Mazurek MA. Atmospheric pressure mass spectrometry: a new analytical chemical characterization method for dissolved organic matter in rainwater. *Environmental Science and Technology* 2003; **37**: 131. DOI: 10.1021/es025848x.
27. Seitzinger SP, Hartnett H, Lauck R, Mazurek M, Minegishi T, Spyres G, Styles R. Molecular level chemical characterization and bioavailability of dissolved organic matter in streamwater using ESI mass spectrometry. *Limnology and Oceanography* 2005; **50**: 1.
28. McIntyre C, McRae C, Jardine D, Batts BD. Identification of compound classes in soil and peat fulvic acids as observed by electrospray ionization tandem mass spectrometry. *Rapid Communications in Mass Spectrometry* 2002; **16**: 1604. DOI: 10.1002/rcm.761.
29. Kramer RW, Kujawinski EB, Zang X, Green-Church KB, Jones RB, Freitas MA, Hatcher PG. Studies of the structure of humic substances by electrospray ionization coupled to a quadrupole-time of flight (QQ-TOF) mass spectrometer. *Special Publication – Royal Society of Chemistry 273 (Humic Substances)* 2001; **273**: 95.
30. Plancque G, Amekraz B, Moulin V, Toulhoat P, Moulin C. Molecular structure of fulvic acids by electrospray with quadrupole time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* 2001; **15**: 827. DOI: 10.1002/rcm.307.
31. Minor EC, Eglinton TI, Boon JJ, Olson R. A protocol for the characterization of oceanic particles via flow cytometric sorting and direct temperature-resolved mass spectrometry. *Analytical Chemistry* 1999; **71**: 2003. DOI: 10.1021/ac981052y.
32. Minor EC, Boon JJ, Harvey HR, Mannino A. Estuarine organic matter composition as probed by direct temperature-resolved mass spectrometry and traditional geochemical techniques. *Geochimica Et Cosmochimica Acta* 2001; **65**: 2819. DOI: 10.1016/S0016-7037(01)00643-3.
33. Minor EC, Simjouw JP, Boon JJ, Kerkhoff AE, Horst J. Estuarine/marine UDOM as characterized by size-exclusion chromatography and organic mass spectrometry. *Marine Chemistry* 2002; **78**: 75. DOI: 10.1016/S0304-4203(02)00011-7.
34. Kim S, Kaplan LA, Hatcher PG. Biodegradable dissolved organic matter in a temperate and a tropical stream determined from ultra-high resolution mass spectrometry. *Limnology and Oceanography* 2006; **51**: 1054.
35. Hockaday WC, Grannas AM, Kim S, Hatcher PG. 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* 2006; **37**: 501. DOI: 10.1016/jorggeochem.2005.11.003.
36. Kim S, Kaplan LA, Benner R, Hatcher PG. Hydrogen-deficient molecules in natural riverine water samples – evidence for the existence of black carbon in DOM. *Marine Chemistry* 2004; **92**: 225. DOI: 10.1016/j.marchem.2004.06.042.
37. Kramer RW, Kujawinski EB, Hatcher PG. Identification of black carbon derived structures in a volcanic ash soil humic acid by fourier transform ion cyclotron resonance mass spectrometry. *Environmental Science and Technology* 2004; **38**: 3387. DOI: 10.1021/es030124m.
38. Kujawinski EB, Del Vecchio R, Blough NV, Klein GC, Marshall AG. Probing molecular-level transformations of dissolved organic matter: insights from electrospray ionization fourier-transform ion cyclotron resonance mass spectrometry. *Marine Chemistry* 2004; **92**: 23. DOI: 10.1016/j.marchem.2004.06.038.
39. Cooper WT, Llewelyn JM, Bennett GL, Salters VJM. Mass spectrometry of natural organic phosphorus. *Talanta* 2005; **66**: 348. DOI: 10.1016/j.talanta.2004.12.028.
40. Koch BP, Witt M, Engbrodt R, Dittmar T, Kattner G. Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization fourier transform ion cyclotron resonance mass spectrometry. *Geochimica Et Cosmochimica Acta* 2005; **69**: 3299. DOI: 10.1016/j.gca.2005.02.027.
41. Llewelyn JM, Landing WM, Marshall AG, Cooper WT. Electrospray ionization fourier transform ion cyclotron resonance mass spectrometry of dissolved organic phosphorus species in a treatment wetland after selective isolation and concentration. *Analytical Chemistry* 2002; **74**: 600. DOI: 10.1021/ac010909f.
42. Stenson AC, Landing WM, Marshall AG, Cooper WT. Exact masses and chemical formulas of individual suwannee river fulvic acids from ultrahigh resolution electrospray ionization fourier transform ion cyclotron resonance mass spectrometry. *Analytical Chemistry* 2003; **75**: 1275. DOI: 10.1021/ac026106p.
43. Marshall AG, Hendrickson CL, Jackson GS. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrometry Reviews* 1998; **17**: 1. DOI:

10.1002/(SICI)1098−2787(1998)17 : 1<1::AID-MAS1>3.0.CO; 2-K.
44. Hannis JC, Muddiman DC. A dual electrospray ionization source combined with hexapole accumulation to achieve high mass accuracy of biopolymers in fourier transform ion cyclotron resonance mass spectrometry. *Journal of the American Society for Mass Spectrometry* 2000; **11**: 876. DOI: 10.1016/S1044-0305(00)00160-4.
45. Kim S, Simpson AJ, Kujawinski EB, Freitas MA, Hatcher PG. High resolution electrospray ionization mass spectrometry and 2D solution NMR for the analysis of DOM extracted by c_{18} solid phase disk. *Organic Geochemistry* 2003b; **34**: 1325. DOI: 10.1016/S0146.6380(03)00101-3.
46. Van Krevelen D. Graphical – statistical method for the study of structure and reaction process of coal. *Fuel* 1950; **29**: 269.
47. Hertkorn N, Benner R, Frommberger M, Schmitt-Kopplin P, Witt M, Kaiser K, Kettrup A, Hedges JI. Characterization of a major refractory component of marine dissolved organic matter. *Geochimica Et Cosmochimica Acta* 2006; **70**: 2990. DOI: 10.1016/j.gca.2006.03.021.
48. Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG. Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra. *Analytical Chemistry* 2001; **73**: 4676. DOI: 10.1021/ac010560w.
49. Qian K, Rodgers RP, Hendrickson CL, Emmett MR, Marshall AG. Reading chemical fine print: resolution and identification of 3000 nitrogen – containing aromatic compounds from a single electrospray ionization fourier transform ion cyclotron resonance mass spectrum of heavy petroleum crude oil. *Energy and Fuels* 2001; **15**: 492. DOI: 10.1021/ef000255y.
50. Knicker H, Hatcher PG. Survival of protein in an organic-rich sediment: possible protection by encapsulation in organic matter. *Naturwissenschaften* 1997; **84**: 231. DOI: 10.1007/s001140050384.
51. Knicker H, Hatcher PG. Sequestration of organic nitrogen in the sapropel from Mangrove Lake, Bermuda. *Organic Geochemistry* 2001; **32**: 733. DOI: 10.1016/S0146-6380(01)00005-5.
52. Kendrick E. A mass scale based on CH2 = 14.0000 for high resolution mass spectrometry of organic compounds. *Analytical Chemistry* 1963; **35**: 2146. DOI: 10.1021/ac60206a048.
53. Kujawinski EB, Behn MD. Automated analysis of electrospray ionization fourier transform ion cyclotron resonance mass spectra of natural organic matter. *Analytical Chemistry* 2006; **78**: 4363. DOI: 10.1021/ac0600306.
54. Rodgers RP, Blumer EN, Freitas MA, Marshall AG. Jet fuel chemical composition, weathering, and identification as a contaminant at a remediation site, determined by fourier transform ion cyclotron resonance mass spectrometry. *Analytical Chemistry* 1999; **71**: 5171. DOI: 10.1021/ac9904821.
55. Rodgers RP, Blumer EN, Freitas MA, Marshall AG. Complete compositional monitoring of the weathering of transportation fuels based on elemental compositions from fourier transform ion cyclotron resonance mass spectrometry. *Environmental Science and Technology* 2000; **34**: 1671. DOI: 10.1021/es990799i.
56. Little DP, Speir JP, Senko MW, O'Connor PB, McLafferty FW. Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. *Analytical Chemistry* 1994; **66**: 2809. DOI: 10.1021/ac00090a004.
57. Senko MW, Speir JP, McLafferty FW. Collisional activation of large multiply charged ions using fourier transform mass spectrometry. *Analytical Chemistry* 1994; **66**: 2801. DOI: 10.1021/ac00090a003.
58. Hockaday WC. The organic geochemistry of charcoal black carbon in the soils of the University of Michigan biological station. Ph.D. dissertation. The Ohio State University. 2006.