


REVIEW

Open Access



# Deciphering dissolved organic matter by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS): from bulk to fractions and individuals

Yulin Qi<sup>1,2,3</sup>, Qiaorong Xie<sup>1</sup>, Jun-Jian Wang<sup>4</sup>, Ding He<sup>5</sup>, Hongyan Bao<sup>6</sup>, Qing-Long Fu<sup>7</sup>, Sihui Su<sup>1</sup>, Ming Sheng<sup>1</sup>, Si-Liang Li<sup>1,2,3</sup>, Dietrich A. Volmer<sup>8</sup>, Fengchang Wu<sup>9</sup>, Guibin Jiang<sup>10</sup>, Cong-Qiang Liu<sup>1,2,3</sup> and Pingqing Fu<sup>1,2,3\*</sup> 

## Abstract

Exploring the source, transformation pathways, and the fate of natural organic matter (NOM) is critical to understanding the regional/global carbon cycle and carbon budget. The dissolved fraction of NOM, i.e., dissolved organic matter (DOM), is a complex mixture resulting from the transformation of plant, animal and microbial matter and plays a crucial role in many biogeochemical processes at the land-ocean-atmosphere interfaces. The advance of Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) makes the detailed characterization of DOM at the molecular level possible. On the other hand, elucidation of complex DOM sample also presents significant analytical challenges, and these challenges also act as a driving force for the instrumentation and methodology development on FT-ICR MS. This review article has been written to aid those working in biogeochemistry, environmental and atmospheric chemistry, and related areas which investigate elemental cycles and DOM transformations. First, the fundamental theory, historical perspective, and recent advances in the field have been introduced. The detailed molecular characterization of environmental and geological samples continues to present significant analytical challenges, and it also has become a driving force for the development of the instrumentation and experimental methods. These achievements in DOM analysis have had an impact upon the fields of environmental science, geochemistry, and analytical chemistry. Next, varieties of applications of FT-ICR MS have also been described, followed by our view of the future of this technique in earth science research. We believe that this review covers the essential pairing of FT-ICR MS and collectively offers environmental and geochemical scientists a substantial resource for their research.

## Highlights

- Principles and procedures of applying FT-ICR MS to decipher DOM molecules were outlined
- Advances of applying FT-ICR MS to reveal the environmental and biogeochemical behaviors were reviewed
- Novel and complementary FT-ICR MS methodologies for DOM structural elucidation in the environment were introduced

**Keywords:** DOM, FT-ICR MS, Organic aerosol, Soil organic matter, Brown carbon, Biogeochemistry, Environmental science

\*Correspondence: fupingqing@tju.edu.cn

<sup>3</sup> Haihe Laboratory of Sustainable Chemical Transformations, Tianjin, China  
Full list of author information is available at the end of the article

## Graphical abstract



## 1 Introduction

Natural organic matter (NOM) is a complex mixture of degraded natural compounds and plays a crucial role in global carbon cycling. Dissolved organic matter (DOM) refers to the dissolved fraction from bulk NOM, which can pass through a  $0.7 \mu\text{m}$  glass/quartz fiber filter. DOM is actively involved in many biogeochemical processes, e.g., cycling of carbon, binding of metals and microbial growth. They offer building blocks and energy sources for aquatic biota, limit light penetration in surface water and interact with anthropogenic compounds (Opsahl and Benner 1997; Wagner et al. 2015; Wu and Tanoue 2001). Additionally, DOM also leads to issues such as water odor, disinfectant demands, membrane fouling, and the production of carcinogenic disinfection byproducts (Zhou et al. 2022). The elemental composition and molecular reactivities of DOM can reveal its role in biogeochemical processes including origin, storage, transport, decomposition and its release as greenhouse gas (Nebbio and Piccolo 2013; Nelson and Siegel 2013).

Understanding the biogeochemistry of DOM is the prerequisite to the management of ecosystem services provided by soil, glacier, lake, riverine, ocean, permafrost, and atmosphere systems. First of all, soil organic matter (SOM) contains at least three times as much carbon as is estimated in either the atmosphere or terrestrial vegetation (Schmidt et al. 2011), and the leaching of SOM leads to the formation of DOM (Roth et al. 2019; Ye et al. 2020). Such an organic carbon pool is undoubtedly sensitive to any changes from the local environment, and in addition, human activities contribute to the biogeochemistry cycle on SOM and DOM (Wang et al. 2019a; Wang et al. 2021c; Zhang et al. 2020a). It is also important to point out that DOM consists of a large variety of functional groups and surface areas, which serve as a suitable container for metals, salts, nutrients, and organic pollutants. The fluxes of continental DOM together with these adducts have important influences on inland waters, lakes, and marine ecosystems. For instance, lake DOM is mostly derived from terrestrial

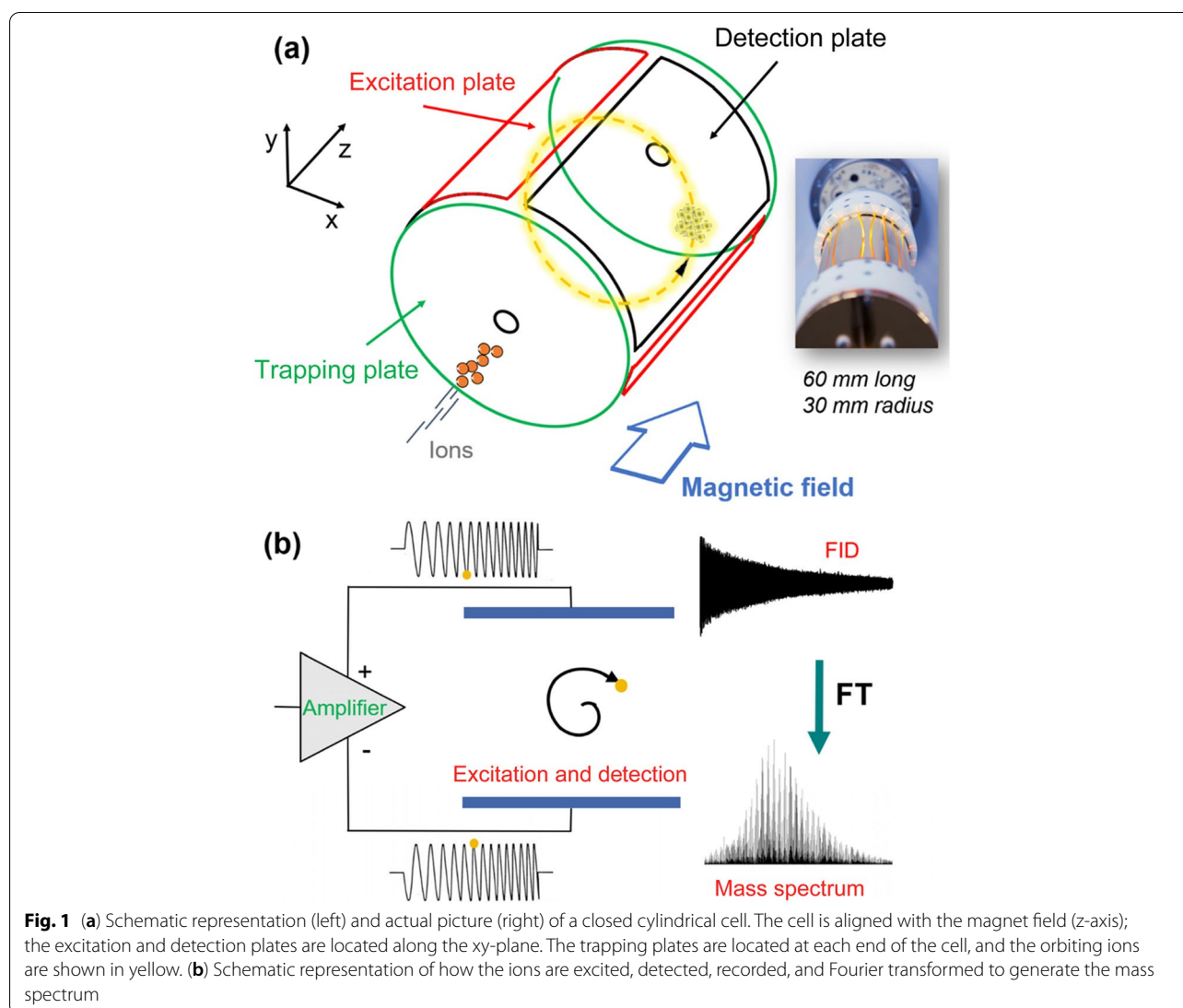
sources (Fu et al. 2006; Wu and Tanoue 2001), and along with freshwater discharge, it can be the determinant of productivity in coastal areas. Meanwhile, the ocean is regarded as the ultimate destination of terrestrial DOM, whose transformation through freshwater systems serves as an important regulator of the quantity of DOM delivered to marine systems (Coble 2007; Tranvik et al. 2009). Last but not least, it has been estimated that the permafrost regions in the Arctic contain 1700 Pg of organic carbon, and in the near future, climate warming in the Arctic will lead to permafrost thaw and cause the release of 41–288 Pg carbon by the year of 2100 (Schoor et al. 2013). The majority of this large amount of ancient DOM will ultimately be transferred to the atmosphere, and give positive and profound feedbacks on climate change (Vonk et al. 2013). Many early studies have been aimed to determine the biogeochemistry cycles of DOM (e.g., Fu et al. 2006), however, achieving a molecular-level understanding of its individual compounds remains the utmost challenge. For this reason, detailed characterization is a crucial step to decipher DOM's chemical role and ultimate fate on a global scale.

Despite recent achievements, a detailed structural characterization of DOM and its variations under different environmental conditions remain largely elusive (Minor et al. 2014). A major issue is that DOM often presents at low concentrations (e.g., 0–10 mg L<sup>-1</sup> in water) compared to the much higher concentration of inorganic salts (approximately 35 g L<sup>-1</sup> in seawater), which can greatly interfere with the analyses (Sandron et al. 2015). The second problem is that a typical DOM sample consists of tens of thousands of compounds. The heterogeneity of their compositions and structures makes it impossible to separate mixtures into individual compounds by liquid chromatography (LC), electrophoresis, or isoelectric focusing (Qi et al. 2022; Sandron et al. 2015).

Bulk analyses using isotopic values, ultraviolet-visible absorption spectrophotometry (Chen et al. 2020) and excitation-emission matrix (EEM) fluorescence spectroscopy with Parallel Factor Analysis (PARAFAC) (Yamashita and Jaffé 2008) can extract information on a large portion of the DOM pool, but do not offer the structural information from the individual components. This limitation remains the same for advanced physical methods such as soft X-ray spectroscopy (Abe et al. 2005), nuclear magnetic resonance (NMR) spectroscopy and fluorescence polarization (Cory and McKnight 2005), which mostly only work well for pure samples. These approaches require isolation and concentration steps prior to analysis, and thus consequently fractionate the DOM pool.

DOM in the atmosphere, water, soil, and sediment is featured by its extraordinary molecular diversities (Chen et al. 2022). Comprehensive chemical and molecular compositional profiles on DOM can be obtained using mass spectrometry (MS) with ultra-high resolution and mass accuracy (MA). MS has been revolutionized by pursuing instruments with increasing mass resolving power (R), making it possible to calculate elemental formulas (C<sub>c</sub>H<sub>h</sub>O<sub>o</sub>N<sub>n</sub>S<sub>s</sub>P<sub>p</sub>... where c, h, o, n, s and p are the numbers of the respective elements) of tens of thousands of compounds simultaneously from accurate mass measurements. Among the different MS instrument types, time-of-flight (TOF), orbitrap and Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) are the three high-resolution mass analyzers. TOF MS can routinely only achieve a broadband R over 10,000 and MA within 5 to 10 ppm, which is much lower than the other two FT instruments (orbitrap and FT-ICR MS) (Boesl 2017). Nevertheless, TOF mass analyzers offer a fast scan rate and no theoretical upper limit for the mass-to-charge ratio (*m/z*). These advantages make it particularly useful for probing large biomolecules such as intact proteins. The orbitrap, a new addition to the family of Fourier transform mass analyzers whose prototype introduced in 1999, has now been widely distributed since its commercial launched in 2004 (Hu et al. 2005). Experiments on a modified Orbitrap Elite instrument showed that R over 1,000,000 could be achieved with appropriate tuning requirements, and such performance allowed the separation of molecule's fine isotopic structure (Denisov et al. 2012). The FT-ICR MS, first introduced in 1974 (Comisarow and Marshall 1974), currently exhibits the highest R and best MA among the various mass analyzers. Lozano et al. (2019) demonstrated a new strategy to acquire a constant ultrahigh R over 3,000,000 across a broad *m/z* range from 260 to 1500 for complex mixture analysis.

Of the wide varieties of MS instruments available, the FT-ICR MS holds the greatest potential for state-of-the-art NOM research, given its inherently ultrahigh R and MA for unequivocal molecule assignments (LeClair et al. 2012; Zhang et al. 2012; Zhang et al. 2020b). For this reason, FT-ICR MS is currently the instrument of choice in geochemistry, environmental chemistry, and atmospheric chemistry for the analysis of organic matter in different environmental matrices at the molecular level, including its properties and reactivity, with sample consumption of only microgram amounts (Qi et al. 2022; Sleighter and Hatcher 2007; Zhang et al. 2020b). However, regardless of the recent achievements, many challenges remain in applying



FT-ICR MS to DOM-related research. This review article summarizes the history, principles, recent advances, analytical strategies, novel applications, and future challenges encountered in the analysis of complex DOM systems.

## 2 Principles of FT-ICR MS

In this section, the most critical principles on FT-ICR MS are briefly introduced; more detailed reviews can be found elsewhere (Qi and O'Connor 2014). Generally, a sample must first be ionized, producing gas-phase ions, which are then analyzed and detected according to their mass-to-charge ratio ( $m/z$ ), generating the mass spectrum by plotting signal intensity against  $m/z$ .

In FT-ICR MS, ions are produced externally from the ion source accumulated in the ion optic system, and transferred into an ion cyclotron cell which is embedded in a

spatially uniform magnetic field ( $B$ ) and an electric field ( $E$ ) (Fig. 1). In the cell, ions with the same  $m/z$  will have the same fundamental cyclotron frequency ( $\omega_c$ ), which is determined by the magnitude of  $B$  and is inversely proportional to the  $m/z$  (Eq. 1a):

Note that the charge ( $q$ ) is calculated by multiplying the number of charges ( $z$ ) by the charge of an electron ( $e$ ); hence,  $m/q$  is related to the measured  $m/z$ . The presence of the electrostatic field offsets the center of ions' cyclotron motion and leads to the reduced cyclotron frequency (Eq. 1b) and the magnetron frequency (Eq. 1c) in the xy-plane.

Cyclotron frequency:

$$\omega_c = \frac{qB}{m} \quad (1a)$$



Reduced cyclotron frequency:

$$\omega_+ = \frac{\omega_c}{2} + \sqrt{\left(\frac{\omega_c}{2}\right)^2 - \frac{\omega_z^2}{2}} \quad (1b)$$

Magnetron frequency:

$$\omega_- = \frac{\omega_c}{2} - \sqrt{\left(\frac{\omega_c}{2}\right)^2 - \frac{\omega_z^2}{2}} \quad (1c)$$

The induced current from the ions can be detected as a function of time and recorded as a composite sum of sinusoidal waves called “free induction decay (FID)”. The FID is Fourier transformed to a frequency domain signal, followed by a mass calibration with respect to the  $m/z$  domain (Qi and O'Connor 2014).

A DOM sample may contain hundreds of thousands of organic species with close mass differences in between and within clusters of nominal mass. For this reason, MA is crucial for the confident assignment of individual elemental formulas within a mass spectrum. The mass error in parts-per-million (ppm) is defined by Eq. 2. Normally, FT-ICR MS can achieve MA at the sub-ppm or even parts-per-billion (ppb) level (Lozano et al. 2019).

$$\text{Mass accuracy} = \frac{m_{\text{experimental}} - m_{\text{theoretical}}}{m_{\text{theoretical}}} \times 10^6 \text{ in ppm} \quad (2)$$

It is also important to point out that MA directly depends on the mass resolving power (R), and a prerequisite for successful analysis is that the signal must be resolved and distinguished from others. For example, a mass split of 0.1 mDa must be resolved in order to assign unique molecular compositions up to 500 Da for DOM species containing C, H, O, N and S. FT-ICR MS is best known for its mass resolving power (Eq. 3), where  $m$  and  $\omega$  are  $m/z$  and the cyclotron frequency for the peak of interest, respectively, and  $\Delta m$  and  $\Delta \omega$  are the peak widths at half maximum (FWHM). DOM studies require R to be over 300,000 to distinguish mass differences of a few millidaltons (Tfaily et al. 2013), for example,  $C_3$  versus  $SH_4$  (3.4 mDa) and  $^{13}C$  versus  $CH$  (4.5 mDa). Moreover, a R of roughly 1,000,000 is required for distinguishing  $^{12}C_4$  and  $^{13}CSH_3$  (1.1 mDa) (Hsu et al. 2011).

$$R = \frac{m}{\Delta m} = \frac{\omega}{\Delta \omega} \quad (3)$$

### 3 History and recent developments

Ultra-high resolution mass spectrometers are indispensable tools for the analysis of complex chemical mixtures. Currently, FT-ICR MS possesses the highest

performance capabilities, has been widely utilized in petroleomics research, and has continued to be a driving force for developing experimental methodologies, data processing, and new instrumentation (Lozano et al. 2020). Advances in petroleomics have been adapted to the study of different complex chemical systems, which then directly push the frontiers of many other fields (Lozano et al. 2020; Qi et al. 2022).

The natural DOM compounds mainly consist of C ( $49.5 \pm 3.3\%$ ), H ( $5.0 \pm 1.0\%$ ), O ( $43.0 \pm 4.1\%$ ), N ( $1.7 \pm 1.0\%$ ), S ( $2.0 \pm 1.3\%$ ) and other elements. Additionally,  $^{13}C$ -NMR data reveal that environmental DOM contains approximately 30% aromatic groups, 23% alkyl groups, 22% carboxyl groups, 20% alkoxy groups, and 5% carbonyl groups (Novak et al. 1992). Nowadays, various analysis methods (e.g., UV, EEM) are available for bulk analysis of DOM samples to estimate their elemental composition and average structure. It is important to note that each method has its unique bias on different analytes. For example, UV and EEM are simple and convenient to characterize the bulk sample, but they are not able to analyze the samples at the molecular level. NMR and MS are modern instruments that can obtain detailed information on individuals. And FT-ICR MS is the technique which has sufficient mass resolving power to separate and accurately assign elemental compositions to individual molecules, thus allowing studies and more robust conclusions about DOM compositions, microbial decomposition pathways, and regional/global cycles. For the above reasons, nowadays, many studies tend to utilize a combination of different methods to show a complementary picture of DOM (Chen et al. 2022; Tfaily et al. 2013; Zhou et al. 2018).

Fievre et al. (1997) utilized FT-ICR MS at the National High Magnetic Field Laboratory (MagLab) to analyze dissolved NOM samples collected from the Suwannee River for the first time, thus creating a new direction and fundamentally changing the science in geochemistry. During the next two decades, more research groups devoted their efforts to this area. In 2018, the 21 Tesla (T) FT-ICR MS at the MagLab was applied to the analysis of DOM samples, demonstrating a resolving power up to 2,700,000 at  $m/z$  400 (Smith et al. 2018).

A total of 571 publications can be searched via “Web of Science” using keywords “FT-ICR MS” and “organic matter” from the year 2011 to 2020 (Fig. 2a). Today, many groups worldwide are working on complex NOM systems, including crude oil, biofuel, bitumen, lignin, sediments, soil, atmospheric aerosols, drinking water, etc. (Fig. 2b). The original works were pioneered by William Cooper (Florida State University) and Alan Marshall (MagLab), and later, by Patrick Hatcher (Old Dominion



#### 4 Extraction of DOM

Atmospheric aerosol and water (e.g., fog water, cloud water, rain and snow) are composed of inorganic constituents, such as sulfate, nitrate, ammonium, sea salt, and a very complex mixture of organic compounds (Andreae and Rosenfeld 2008; Bao et al. 2018; Bruggemann et al. 2020; Chen et al. 2022; Kundu et al. 2012; Su et al. 2021; Wu et al. 2021; Xie et al. 2020; Zhao et al. 2013). Water-soluble organic matter (WSOM) consists of a substantial fraction (20–80%) of particulate organic matter (Sullivan et al. 2004), which is light-absorbing and contributes to global warming (Yue et al. 2022). Moreover, inorganic ions usually cause interferences during the WSOM characterization. Therefore, separation of the WSOM from inorganic ions is crucial for studying the organic compounds in atmospheric aerosol and atmospheric water (Lin et al. 2012). The concentration of dissolved organic carbon in the deep ocean is typically lower than  $50 \mu\text{mol L}^{-1}$ . This very low DOM concentration with large amounts of inorganic salts present in water samples simultaneously causes significant analytical challenges. De-salting, isolation and concentration of the DOM components are the prerequisites for DOM characterization.

Extraction of DOM can typically be achieved by one of four ways: (1) ultrafiltration, (2) direct drying or freeze-drying, (3) reverse osmosis coupled with electrodialysis (RO/ED), or (4) solid-phase extraction (SPE). SPE has been shown to be the “easier and quicker technique” compared to the others (Simjouw et al. 2005). Therefore, it is the most adapted extraction method. In SPE, molecules are retained on an appropriate solid phase while the DOM sample flows through a prepacked column; compounds of interest are then eluted using suitable solvents. Studies show that C-18 and styrene divinylbenzene-based sorbents are the most effective for DOM extraction across a wide range of water types; the recovery of the DOM components depends mainly on the polarity of the molecules (Minor et al. 2014). Dittmar et al. (2008) examined polar to highly polar prepacked polymer and silica-based sorbents to compare differences of extracted DOM molecules from seawater using different SPE cartridges. The authors found that the styrene divinylbenzene polymer was most efficient, as it recovered 62% of DOC as salt-free extracts. Subsequent  $^1\text{H}$ NMR, C/N and  $\delta^{13}\text{C}$  analyses also confirmed that styrene divinylbenzene extracted a more representative proportion of DOM than the other materials; and hence, this extraction protocol has become a widely recognized method for simple and reproducible solid-phase extraction of DOM from water samples. It should be noted that the acidification during the SPE

extraction can impact the molecular composition of DOM in different environmental matrices (Han et al. 2022).

Soil organic matter (SOM) is a category of complex NOM, which is a crucial reservoir for carbon and nutrient biogeochemical cycling in soil-based ecosystems. For SOM characterization, organic compounds must be extracted from the soil with little or no chemical alteration (Bahureksa et al. 2021; Wang et al. 2021b; Zhang et al. 2020a). Tfaily et al. (2015) investigated a broad range of chemically diverse soil types using a combination of solvents with varying polarity. The FT-ICR MS analysis of the extracted SOM indicated that hexane was suitable for lipid-like compounds with very low O/C ratios, while water was selective for biological significant molecules such as carbohydrates and amino-sugars with high O/C ratios. In addition, acetonitrile extracted more compounds in the tannin and lignin-like regions with  $\text{O/C} > 0.5$ ; methanol was more suitable for the compounds with  $\text{O/C} < 0.5$ . However, given that no extraction approach can fully recover the DOM components from an intricate environmental sample, a thorough methodology selection should be conducted based on the sample conditions before the FT-ICR MS analysis.

#### 5 Ionization techniques on DOM coverage

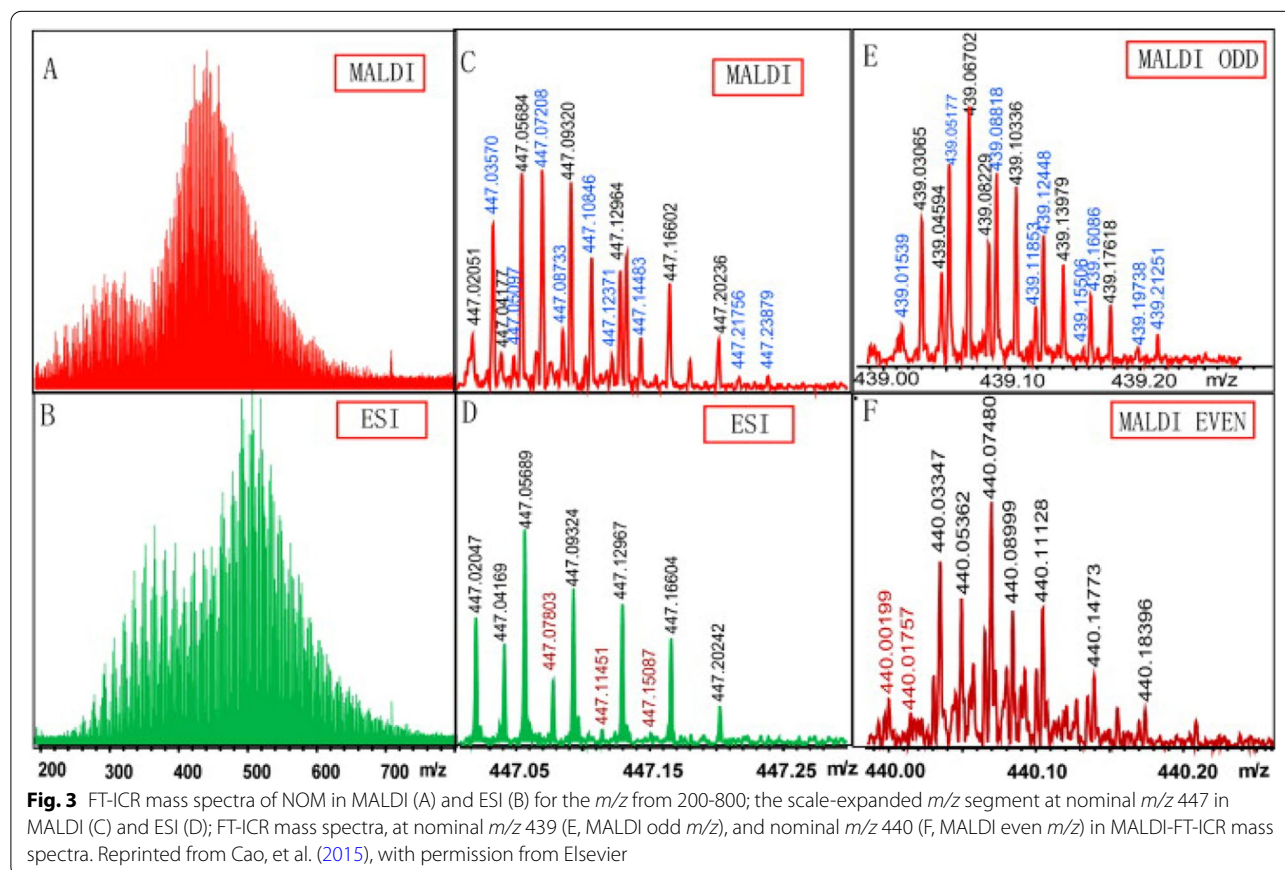
The ionization selectivity for DOM is mainly determined by two fundamental groups in a molecule: polar and non-polar groups. MS was not applied for the DOM analysis until the invention of the electrospray ionization (ESI) source in the late 1990s. Such a technique ionizes compounds from the aqueous solution and produces protonated or deprotonated molecular ions ( $[\text{M}+\text{H}]^+$  or  $[\text{M}-\text{H}]^-$ ) for the subsequent MS analysis (Fenn et al. 1990). Predominantly, negative electrospray ionization (ESI-) is utilized for the MS study of DOM due to its high selectivity for polar acidic functional groups, requiring only a very small sample amount for analysis. Nevertheless, the ESI spectra are often cluttered by salt adducts, while compounds containing sulfur (S) or nitrogen (N) are often lost. In addition, size exclusion chromatography also shows that higher molecular weight and UV-active compounds are often excluded from ESI spectra (Hawkes et al. 2019).

During ESI, the sample solution forms charged droplets via a high voltage. After the droplet evaporates, the charge is concentrated and once the charge exceeds the Rayleigh limit, the droplet dissociates and generates ions (Fenn et al. 1990). According to this mechanism, more polar components are favored by ESI, with primary polar species formed in positive mode and acidic polar species in negative mode.



The carboxyl group is probably the most important moiety in DOM molecules, as it controls DOM's acidity, buffering activity, water retention and metal-binding strength, etc. The DOM pool consists of a large portion of carboxyl-rich alicyclic molecules (CRAM), including both aromatic and aliphatic carboxylic functions, and this makes ESI- the routine method for MS in geoscience studies. However, ion suppression is observed when measuring complex DOM samples, and the ion intensity is often reduced due to space charge effects. Moreover, DOM compounds vary in heteroatom content, functional groups, molecular size and stereochemical conformation etc., which cause charge competitions between ions during the ESI process. For all the reasons above, there is no universal ionization technique or MS parameter setting that will detect all of the DOM components with equal efficiency, so the mass spectra of a DOM sample measured by various ion sources often differ dramatically in the spectral appearance (Qi et al. 2020). As a result, a proper selection of the ionization conditions or even a combination of ionization methods should be considered for comprehensive characterization of the environmental sample.

For research purposes, various atmospheric pressure ionization (API) sources such as atmospheric pressure photoionization (APPI), atmospheric pressure chemical ionization (APCI), laser desorption/ionization (LDI) and matrix-assisted laser desorption/ionization (MALDI), have been investigated for the study of DOM and related substances. For example, in APPI, the sample solution vaporizes and absorbs photons emitted by a discharge lamp to form ions; therefore, APPI ionizes less polar compounds as compared to ESI. D'Andrilli et al. (2010b) compared the FT-ICR MS analysis of marine DOM. It was found that APPI produced ions with higher carbon unsaturation, whereas ESI generated more oxygenated ions. Moreover, APPI could also ionize polar species and was more tolerant to matrix effect from solvents and salts, thus allowing molecular characterization over different degrees of hydrogen saturation and oxygenation. Podgorski et al. (2012) also showed that an N-containing ion species with the same molecular formula generated by APPI exhibited a 13 times higher S/N than that by ESI. Similar results were demonstrated by Qi et al. (2020), who systematically evaluated the performance of ESI, APPI and APCI using a commercial lignin standard. The authors showed that the number of heteroatoms (e.g., N, S, P) in the lignin molecule





significantly increased its chemical diversity. For example, S is more likely to form thiol groups as observed by ESI, while the S-based thiophenic species linked to the aromatic core are uniquely detected by APPI. In 2018, paper-based spray ionization was also adapted for the analysis of DOM, which showed more tolerance to salt contamination than ESI (Kim et al. 2018).

Compared to API, matrix-assisted laser desorption ionization (MALDI) is less prevalent in DOM studies, primarily because of its lower ionization efficiency using conventional matrices (Qi and Volmer 2019), which inhibit the reproducible acquisition of high-quality mass spectra. Nevertheless, MALDI measures solid-phase sample within milliseconds, and hence, it can be directly applied as a shotgun method to determine the most abundant compounds in various environmental samples. For instance, Solihat et al. (2018) demonstrated a proof-of-concept experiment for using this method to analyze soil samples. Only 500  $\mu\text{g}$  of unprocessed soil powder was fixed on a metal plate using double-sided adhesive tape and analyzed directly using FT-ICR MS. Such an approach avoided the extensive extraction and sample desalting steps that other API techniques required. Cao et al. (2015) also compared DOM analysis by MALDI and ESI, demonstrating the complementary nature of the two techniques, often, molecular information invisible by ESI was observed by MALDI. Interestingly, although most peaks at even  $m/z$  numbers have identical molecular formulas compared to the corresponding peaks at odd  $m/z$ , an average difference of  $\sim 1.007825$  Da (corresponding to the atomic mass of hydrogen) was observed between even and odd  $m/z$  values (Fig. 3). Therefore, the authors speculated that each peak at even  $m/z$  having the same molecular formula as that at an odd  $m/z$  was actually a molecular ion  $[M+e]^-$  (the mass of an electron ( $e$ ) is 0.000549 Da). Moreover, the related peak at odd  $m/z$  was a deprotonated molecule  $[M-H]^-$  (the mass of a hydrogen proton is 1.007276 Da). These radical anions were successfully distinguished by ultra-high resolution of FT-ICR MS (Cao et al. 2015).

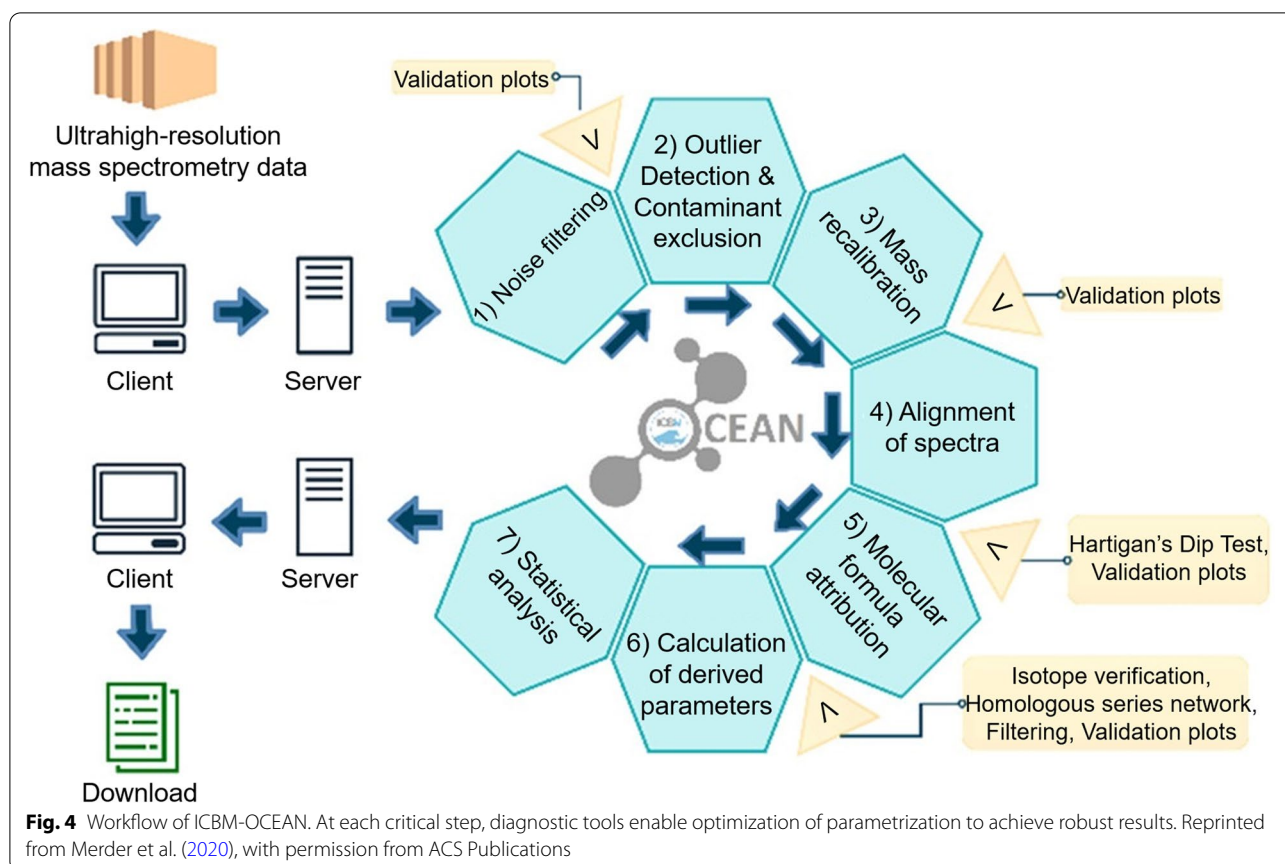
Overall, different ionization methods exhibit different selectivities for DOM molecules. For a comprehensive study of DOM, the choice of method should be tailored to meet the particular measurement purposes, cover the desired compound classes, and prevent erroneous interpretations from a single ionization source.

## 6 Data visualization and processing

Considering the immense complexity of DOM mass spectra, simplification and classification are necessary to process the large datasets and extract the relevant information efficiently from the mass spectra. Numerous graphical and statistical methods have been developed

to study DOM to visualize data, analyze samples, enable fingerprinting, and facilitate comparisons between datasets. In general, the van Krevelen plot is most widely utilized, which helps to sort the same type of compounds in specific regions and makes it straightforward to visualize possible links between molecules (Kim et al. 2003). Bar charts such as oxygen, carbon number and double bond equivalent (Tose et al. 2018) distribution have been plotted to examine the chemical features of DOM. The Kendrick mass defect (KMD) rescales the molecules'  $m/z$  units according to their homologous structural units (e.g., 14.01565 for  $\text{CH}_2$ ) and allows the alignment of thousands of compounds with the same functional groups (Qi et al. 2016). The above methods are being continuously improved to extract hidden information from the FT-ICR MS data. For example, the modified aromaticity index ( $\text{AI}_{\text{mod}}$ ) is calculated to reflect the carbon unsaturated bond density in a given elemental formula. Such a parameter is typically used in combination with a van Krevelen plot for data visualization; compounds with  $\text{AI}_{\text{mod}} \geq 0.5$  are usually identified as aromatic compounds,  $\geq 0.67$  are unambiguous indicators of polycyclic condensed aromatics, while  $< 0.5$  corresponds to non-aromatic compounds (Koch and Dittmar 2006). Rivas-Ubach et al. (2018) proposed the multidimensional stoichiometric compound classification (MSCC) to classify the assigned DOM molecules into lipid-, peptide-, aminosugar-, carbohydrate-, nucleotide-like compounds according to the C/H/O/N/P stoichiometric ratios. D'Andrilli et al. (2015) also calculated a molecular lability boundary (MLB) based on the FT-ICR MS data to complement the van Krevelen diagrams. By dividing the data into more and less labile constituents, DOM compounds above the MLB at  $\text{H/C} \geq 1.5$  correspond to more labile materials, while those below the MLB,  $\text{H/C} < 1.5$ , exhibit less labile, more recalcitrant character. The application of the MLB also offers a lability threshold to compare data from other FT-ICR MS that analyze DOM samples from around the world, e.g., DOM in glacial was found to be most labile, followed by that in marine and freshwater systems. Mann et al. (2015) proposed a CHO index based on C, H and O in the molecule to codify SOM components according to their observed degradation potentials: compounds with an index between -1 and 0 showed high degradation potential, while similar stoichiometries in the base-soluble component did not. In addition, the CHO index allows for direct identification of the distribution of heteroatoms in the soil sample when compared with the van Krevelen diagram.

Due to the ultra-complexity of FT-ICR MS data, mathematical statistics are becoming commonplace; multi-variate analyses can translate raw DOM data with



different sizes, structures, and sources into a model. For example, principal component analysis (PCA), redundancy analysis (RDA) and Spearman's correlation analysis have widely been applied to petroleomics and environmental samples to indicate the variability and correlate molecular characterization between samples (Li et al. 2018). More recently, Hur et al. (2018) introduced a scatterplot of univariate analysis, so called "volcano plot", to handle the DOM from the petroleomics sample. Such plots were generated by plotting p-value fold change. The differences between two large datasets can therefore be clearly visualized in a statistically significant way and the filtered peaks are plotted using traditional DBE versus carbon number and van Krevelen plots for further analysis.

## 7 Software developments

Mass spectral processing and characterization of complex DOM compounds including signal-to-noise ratio (S/N) filtering, mass recalibration, chemical formula assimilation, isotopic validations, and graphic visualization, etc., are nowadays frequently performed via freely available or commercial computer programs.

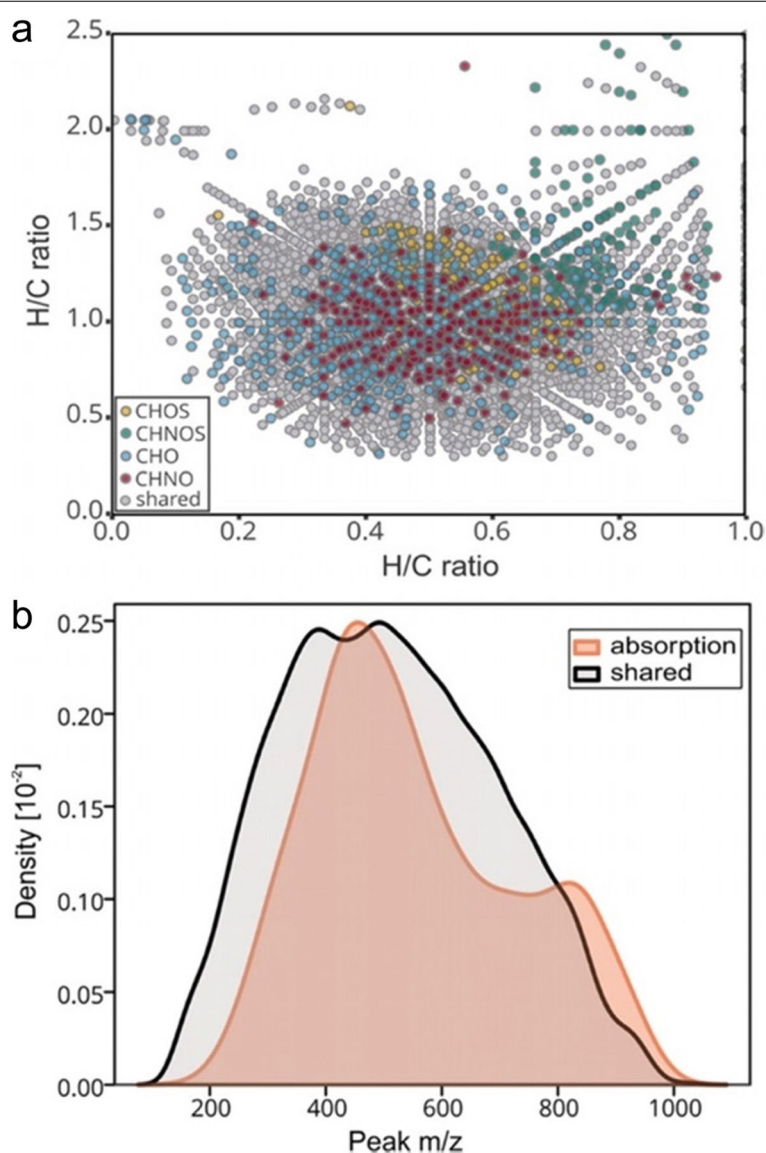
The freely available Formularity software developed by Andrey Liyu and Nikola Tolić at Pacific Northwest National Laboratory, USA, is capable of assigning chemical formulas to FT-ICR MS spectra of DOM by comparing measured  $m/z$  values with theoretical  $m/z$  values of pre-constructed formula library and selecting the optimum formula among multiple ambiguous solutions based on the minimum number of non-oxygen heteroatoms and homologous inspection (Tolić et al. 2017). Kew et al. (2017) published an MS visualization program by proposing an interactive web browser-based graphical form. This tool provides centroid mass spectra, van Krevelen diagram, DBE versus carbon number plots and  $AI_{mod}$  versus carbon number plots, offering a powerful visualization of data from DOM mass spectra. However, the entire spectral processing must be performed on the user's computer using command lines, thus requiring sufficient computer knowledge of the user. Leefmann et al. (2019) also proposed a fully web browser-based tool called Ultra-MassExplorer (UME). The website offers a complete data evaluation process, from the calculation of elemental compositions to data visualization. The mass spectra can be uploaded in ASCII or CSV formats, consisting of  $m/z$  signals and intensities. The processed data can be

reviewed by interactive visualizations including 2D and 3D van Krevelen plots. However, some limitations should also be mentioned for these algorithms. For example, Formularity requires input of mass spectral data manually and provides relatively poor performance for peaks with  $m/z$  over 500. Similarly, multiple formulas can be assigned for a single peak by UME, and therefore, a further filtration is required to select the formula of best fit.

Merder and Dittmar recently launched a server-based tool called “ICBM-OCEAN (Institute for Chemistry and Biology of the Marine Environment,

Oldenburg–complex molecular mixtures, evaluation & analysis)”. This program integrates published novel approaches, including removal of noise, sidebands and contaminants, alignment of spectra from different samples, isotope ratio filter linear quantile regression filter, and fundamental statistical analysis (Fig. 4) in order to achieve standardized analysis of DOM acquired from FT-ICR MS (Merder et al. 2020; Merder et al. 2019; Riedel and Dittmar 2014).

Similarly, the MFAssignR Code based on R language could estimate noise, filter  $^{13}\text{C}$  and  $^{34}\text{S}$  polyisotopic



**Fig. 5** (a) Van Krevelen diagram showing the compound class of new molecular formulas assigned to shared peaks in triplicate measurements of the absorption mode (peaks shared between modes are shown in gray); (b)  $m/z$  density distribution of the peaks shared between the modes and additionally assigned in the absorption mode. Reprinted from Da Silva et al. (2020), with permission from the ACS Publications

masses, recalibrate masses, and assign molecular formulas for Suwannee River and aerosol DOM samples (Schum et al. 2020). The free TRFu Code was compatible to the negative FT-ICR MS spectra of different NOM types (e.g., freshwater, marine, soil, atmospheric aerosol, biochar) with the formula assignment accuracy up to 94% (Fu et al. 2020a). The TRFu Code was further updated to the NOMDBP Code and FTMSDeu Code to simultaneously assign chemical formulas to NOM and organohalogen molecules (Fu et al. 2020b; Fu et al. 2022). In addition to  $^{13}\text{C}$ ,  $^{37}\text{Cl}$ , and  $^{81}\text{Br}$  isotopic patterns, three filter rules and an empirical mass distribution rule of iodinated molecules have been proposed to improve the formula assignment accuracy for organohalogens. The FTMSDeu Code is also compatible to deuterium labeled FT-ICR MS spectra, providing a powerful tool in elucidating liable hydrogen atoms (mainly hydrogen in carboxyl and hydroxyl functional groups) and non-liable hydrogen atoms (such as hydrogen in carbonyl or ether group) in DOM molecules.

Regarding commercial software, PetroOrg (by the MagLab) and Composer (by Sierra Analytics, Inc.) are highlighted, which were developed by experts in the analysis of crude oil and related materials. However, both software programs were initially developed to meet the demands of the petroleomics community and thus parameters should be more carefully optimized to adapt them to the organic geochemistry community.

While FT-ICR MS is a powerful technique to resolve the complexity of DOM mixtures at the molecular level, even the ultra-high resolution of FT-ICR MS may result in ambiguous or false formula assignments to peaks in a DOM mass spectrum, especially at the high  $m/z$  region. Phase correction is a data processing method that increases the mass resolving power of any Fourier transform instruments by a factor up to two and the S/N by square root of 2 compared to the conventional magnitude mode (Qi et al. 2012). Although software for absorption mode processing is available for both FT-ICR and Orbitrap MS, it has not been widely applied in DOM studies yet. For this purpose, Da Silva et al. (2020) tested the absorption mode for analyzing a reference sample -- Suwannee River natural organic matter (SRNOM), using analytical figures of merit including precision, accuracy, and reproducibility. The result showed that the additional effort for processing the absorption mode mass spectra achieved three times higher mass resolving power and a pronounced improvement for molecules containing nitrogen and sulfur (Fig. 5).

Overall, untargeted molecular analyses of complex mixtures are relevant for many research fields, including proteomics, petroleomics and geochemistry. The demand for more efficient data processing, visualization,

fingerprinting and comparisons between samples is snowballing. Open scripts, online tools and commercial software for DOM data processing will undoubtedly further flourish in the community and receive more attentions in the future.

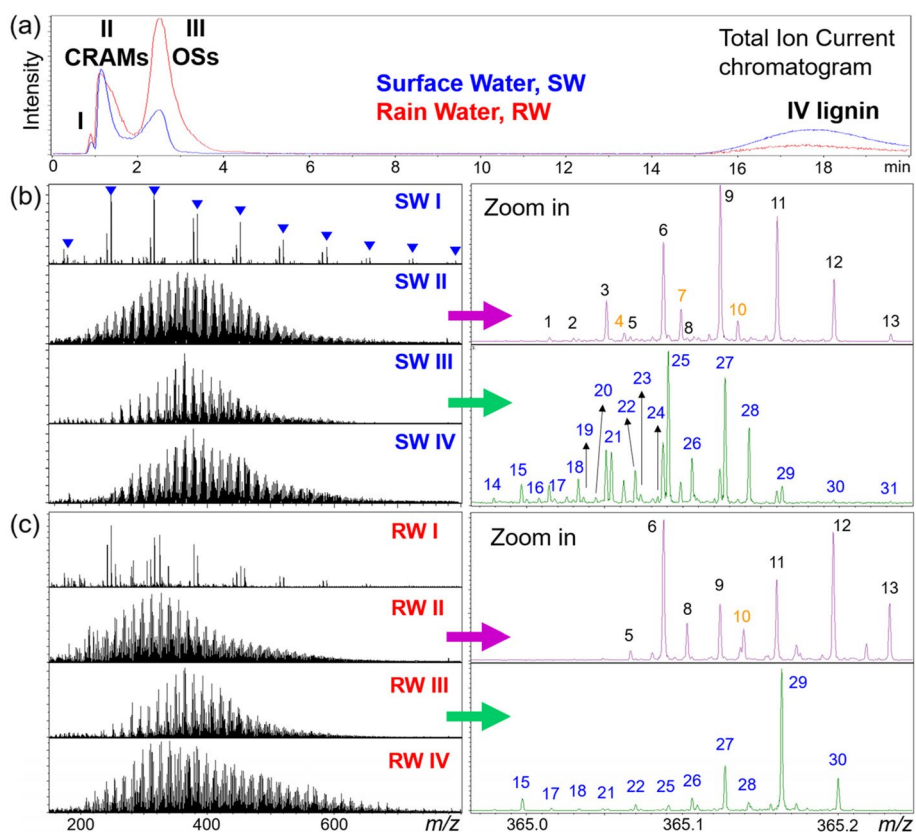
## 8 Paradigm in geochemistry and environmental science

Determining the detailed composition of DOM in different environments is the prerequisite to understand its impact on water quality, pollutant bioavailability, as well as its long-term effect on microbial/anthropogenic activities. During the last two decades, FT-ICR MS has become a well-established analytical tool in the community for detailed DOM analyses. In the limited space of this review, we were not meant to be comprehensive but rather highlight some of the important milestones in FT-ICR MS application originating from around the world to review the paradigm of FT-ICR MS in geochemistry studies.

In the early 2000s, the community adopted KMD plots to visualize the mass spectra. Such a plot allows compounds with the same functional groups or structural units to be aligned into a compact visual space, without calculation of elemental formula. For example, the molecular mass of each molecule can be normalized to the  $\text{CH}_2$  group, to generate a horizontal row of the points in the plot that representing a homologous series of molecules separated by 14.0000 Da. By contrast, in the vertical row, higher KMD values indicate that these molecules are more saturated (Stenson et al. 2003). Later, Kim et al. (2003) introduced the van Krevelen plot to sort the compounds according to their O/C and H/C ratios. The diagram makes it straightforward to visualize the components within the same category as well as their possible links. For example, Mopper et al. (2007) utilized the diagram to map the humic components (consisting of polar, carbonyl groups) collected from the reversed-phase LC, and it was found that fractions with higher O/C ratios eluted off faster than the lower ones.

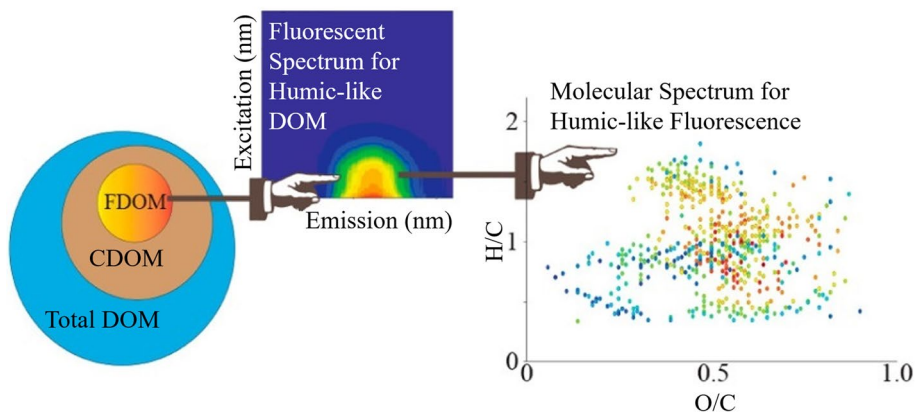
Although the van Krevelen plot has become an attractive and necessary step to provide valuable insight into the DOM data, its drawbacks should be kept in mind. Firstly, the assignment of elemental formulas is required for the calculation of O/C and H/C ratios. Secondly, the information on molecular mass is lost in such a two-dimensional diagram. Thirdly, the boundary of compound classes is not notable in the plot and the fractions tend to be centralized in the lignins and tannins territory. Lastly, the calculated elemental formula represents molecule's composition only, with no connection to the hundreds or thousands of possible structural isomers. For example, a recent study using





**Fig. 6** (a) Total ion current chromatogram for surface water (SW, blue) and rainwater (RW, red) samples by LC-FT-ICR MS with four different fractions labeled. (b) Averaged mass spectra of SW for the four fractions (left) and mass scale expansion of the region of  $m/z$  364.96–365.25 of fractions SW II and SW III. The nitrogen-containing compounds are labeled in yellow, and the sulfur-containing compounds (mainly OSs) are labeled in blue. (c) Averaged mass spectra of RW for the four fractions and mass scale expansion of the region of  $m/z$  364.96–365.25 of fractions RW II and RW III. Reprinted from Qi et al. (2021), with permission from the ACS Publications

**What molecular families track fluorescence in natural waters?**



**Fig. 7** The van Krevelen plot shows the molecular families associated with PARAFAC components and data points are coloured by molecular weight. Reprinted from Stubbins et al. (2014), with permission from the ACS Publications

TIMS-FT-ICR MS demonstrated that approximately 6 to 10 isomers could be assigned for each elemental formula (Leyva et al. 2019). Therefore, the ratio of elements is merely a reference to identify the compound classes. For the above limitations, additional criteria such as CHO index,  $AI_{mod}$ , MLB are further applied to complement the van Krevelen plot.

While FT-ICR MS is a powerful instrument for the analysis of complex DOM samples, the combination with another analytical technique further increases separation power. As early as 1997, Fievre et al. (1997) coupled LC with FT-ICR MS, and demonstrated that individual structural information on fulvic acid mixture could be obtained. The reason for the improved performance is the increased ionization efficiency of the compounds of interest, as they are no longer in the presence of various other types of constituents. Recently, Kim et al. (2019) demonstrated that the separation of lignin- and tannin-type compounds from NOM can be achieved via LC coupled with FT-ICR MS. In this work, they were able to perform both quantitative comparison and molecular level investigation on NOM with a sample of 2  $\mu$ g. Additionally, Qi et al. (2021) utilized LC-FT-ICR MS for online analysis of DOM in river- and rain-water. With sophisticated instrumental optimization, different portions of metal salts, carboxyl-rich alicyclic molecules, organosulfates (OSs), and lignin-like compounds were successfully fractionated within one LC cycle of 20 min (Fig. 6).

The light-absorbing fraction of DOM is called chromophoric DOM (CDOM), which directly affects the optical properties of the aquatic ecosystems and protects aquatic organisms from UV radiation. It should be noted that the light-absorbing fraction of organic aerosol in the atmosphere is generally called brown carbon (BrC) (Hu et al. 2021; Laskin et al. 2015; Li et al. 2020; Yue et al. 2022). When excited by photons with specific energy, a sub-fraction of CDOM emits fluorescence; this portion of DOM is called fluorescent DOM (FDOM). Due to these properties, UV and fluorescence spectroscopies have been widely utilized to investigate the aromaticity, sources, reactivities and diagenetic states of DOM (Coble 2007). In addition, such spectroscopic tools can provide quantitative information, which is considered a major disadvantage of FT-ICR MS for DOM analysis (Li and Hur 2017) (Ge et al. 2022). Tremblay et al. (2007) coupled UV and EEM with FT-ICR MS to trace the transport and changes of DOM from mangrove porewaters in a Brazilian estuary. Subsequently, the peak picking algorithm to define fluorescent fractions in EEM was improved by the statistical fitting approach, PARAFAC, which can break down the fluorescence signal into underlying individual fluorescent components with

specific excitation and emission spectra (Stedmon and Bro 2008).

Stubbins et al. (2014) analyzed river water samples from boreal Québec, Canada, using the EEM/PARAFAC approach and FT-ICR MS to define the molecular signatures associated with EEM (Fig. 7). It was found that 39% of the molecular formulas identified by FT-ICR MS were related to the PARAFAC components, showing that the coupling of FT-ICR MS and EEM/PARAFAC offered information to the biogeochemical cycling of a much larger proportion of the DOM pool.

The negative-mode ESI is the most applied ionization method for DOM analysis, making FT-ICR MS biased for hydro-soluble compounds. By contrast, GC-MS is suitable for volatile compounds, which are generally organo-soluble (Fu et al. 2012; Tang et al. 2020; Zhang et al. 2020b). In addition, pyrolysis (Py) can be coupled with the GC-MS to crack the non-volatile DOM components into small volatile molecules at a high temperature (e.g., 800 °C), and therefore, more precise information can be quantitatively obtained by GC-MS, with very small amounts of sample. For instance, Py-GC-MS is a classical method to determine the source, degradation state, and specific functional units for terrestrial lignin (Goñi and Hedges 1992; Hedges et al. 1982; Louchouart et al. 2010). Nevertheless, such a method derivatizes and liberates a fraction of the monomers for analysis, and hence, only provides partial data due to the specificity of the various treatments. So, by the integration of GC-MS, or two-dimensional GC-MS with FT-ICR MS, both organo- and hydro-soluble DOM can be well characterized in detail (Drosos et al. 2017; Xu et al. 2021).

Fingerprinting of DOM and determining its fate in various environments is crucial to understanding the global carbon cycle. Oceans are a large and important natural environment. Marine DOM represents the largest organic carbon reservoir on Earth, playing an important role in many biogeochemical processes (Hedges et al. 2000). Usually, studies in the field of marine science span several sampling sites, and FT-ICR MS analysis is utilized to reveal the sources, similarities, and differences of DOM from site to site; formation and degradation processes; environmental impacts and potential reactive nature. Several representative publications in this field deserve special mention, including Hertkorn' and Dittmar's work on the Pacific Ocean (Dittmar and Koch 2006; Hertkorn et al. 2006), Koch's work on Antarctic bottom waters (Koch et al. 2008) and D'Andrilli's work on the Weddell Sea (D'Andrilli et al. 2010b). The above paradigm is further expanded to investigate the DOM flux from fluvial systems into oceans and the anthropogenic influences from watersheds (He et al.

2019; He et al. 2020; Pang et al. 2021; Roebuck Jr et al. 2019; Sleighter and Hatcher 2008; Wang et al. 2021a; Wang et al. 2019b). An underlying problem in these works is to distinguish the composition of marine DOM from terrestrial DOM. Koch et al. (2005) established the groundwork by comparing the molecular composition of a marine DOM from the Weddell Sea with terrestrial DOM from the coast of Northern Brazil and found surprisingly, that one-third of the molecular formulas were observed in both samples. The authors assumed that the similarities were due to biotic and photochemical processes, independent of the environmental conditions where the DOM samples came from. Later on, this hypothesis was also confirmed by another study (Zark and Dittmar 2018).

Peatland is a major reservoir of organic carbon, which is climate-sensitive and can be greatly affected by ground temperature, warming air, soil respiration and microorganisms. Peatland DOM mainly consists of cellulose and lignin-derived phenols from plants (Benner et al. 1984). It is known that warm climate enhances microbial productivity and makes the peatland DOM more vulnerable to degradation. Such a phenomenon causes a positive feedback to climate change by turning the carbon reservoirs into emitters of greenhouse gases ( $\text{CO}_2$  and  $\text{CH}_4$ ). However, due to the complex geochemical conditions, it is difficult to characterize the decomposition pathways of DOM. For example, the DOM composition in peatland changes with depth and varies between alkaline fens and more acidic

bogs. FT-ICR MS was utilized to reveal compositional differences of peatland DOM as well as their reactivities. D'Andrilli et al. (2010a) reported the DOM in fen and bog porewaters from the Glacial Lake Agassiz Peatlands of northern Minnesota, USA and concluded that hydrogenotrophic methanogenesis is related to the less reactive NOM in bogs, while moister conditions may benefit sedge-dominated fens and increase the pool of more labile unsaturated and aromatic compounds. A subsequent study indicated that oxygen shedding and DOM reduction are important processes during peat decomposition, which may explain the increased  $\text{CO}_2$  concentrations relative to  $\text{CH}_4$  in peatlands and can serve as an indicator for climate feedbacks due to the higher global warming potential of  $\text{CH}_4$  compared to  $\text{CO}_2$  (Tfaily et al. 2013).

The above applications of FT-ICR MS coupled with other analytical instruments have offered an appreciation of the extent to which specific DOM compound in soils, surface waters, glaciers, sediments, and atmosphere could play important roles in a variety of biogeochemical processes. With the development of modern analytical techniques, the fate of DOM in these environments has also become apparent and begun to attract many new practitioners.

When viewing these paradigmatic studies, it is obvious that the van Krevelen plot has been adapted as a routine strategy to visualize the complex FT-ICR MS datasets. Nevertheless, it is also important to point out that this approach fits the compounds into a two-dimensional

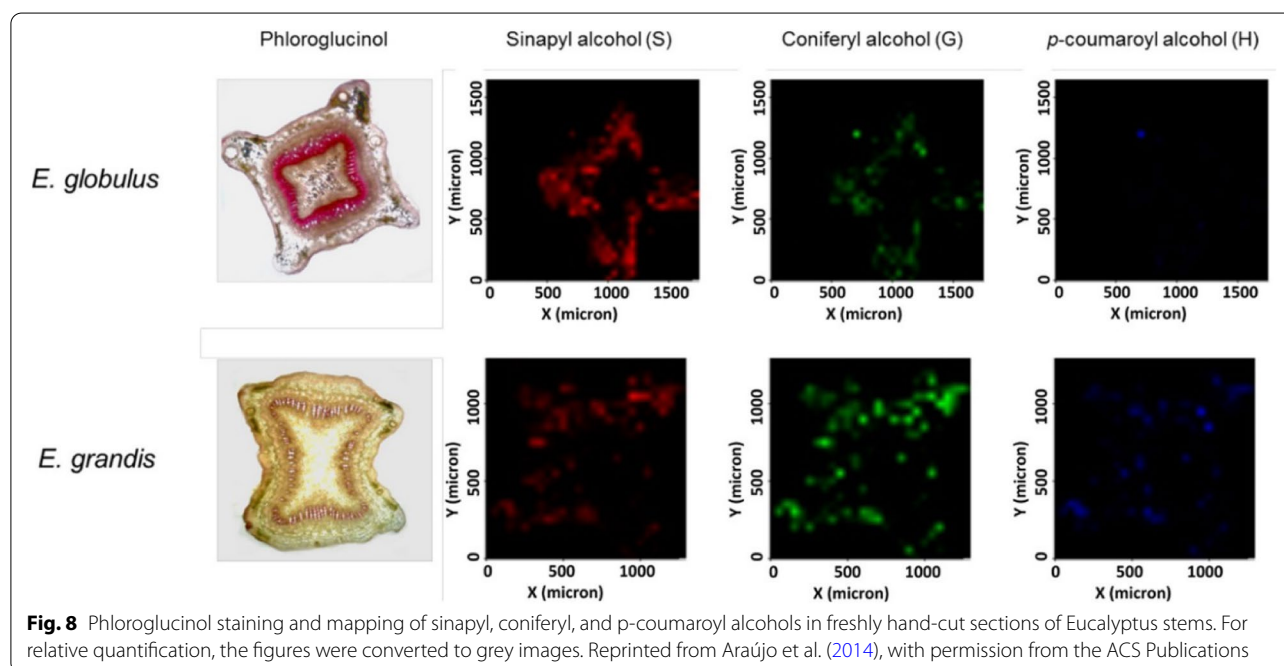


diagram according to the O/C and H/C ratios, so that information on the molecular weight is lost. More importantly, each data point in the plot may represent a large number of distinct structural isomers! To overcome these issues, complementary data handling and analytical tools are increasingly applied in the field to better explore the geochemical information hidden in the DOM mixtures.

Last but not least, data processing such as PCA and RDA utilized the abundances of the ion signals; however, the signal abundances in the FT-ICR MS measurements are greatly affected by the experimental parameters of the instrument. Therefore, the parameter setting of FT-ICR MS should keep constant for applying the statistical analysis to ensure statistical confidence and accurately correlate the features between DOM and environment (Hawkes et al. 2020).

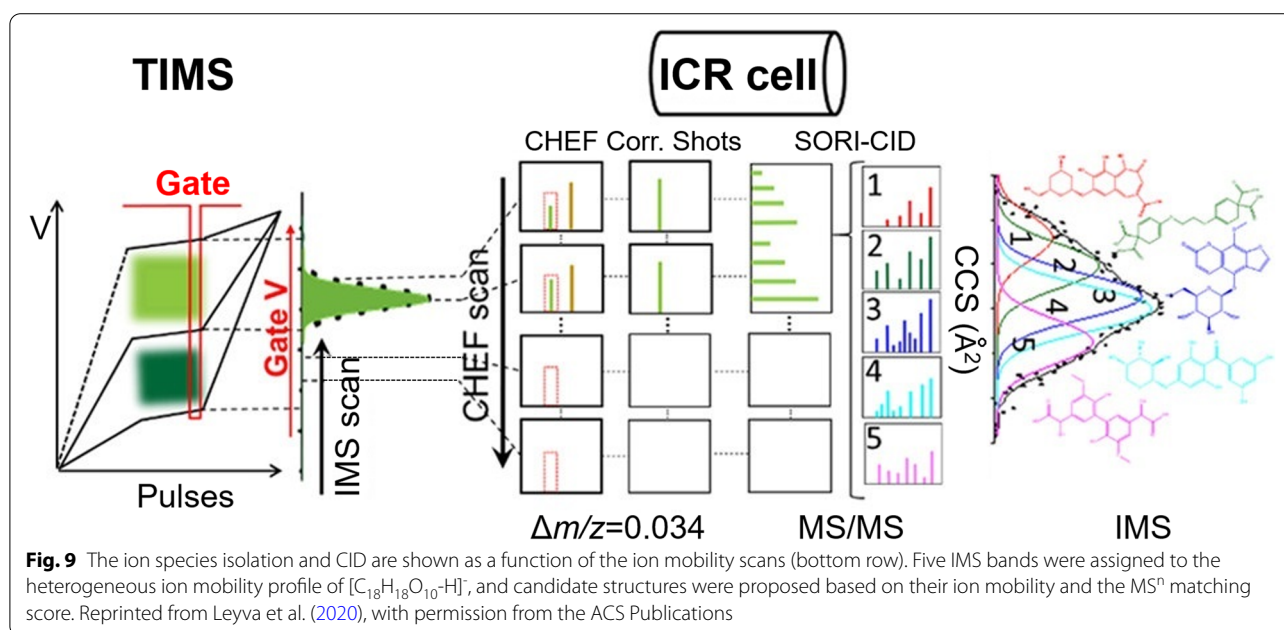
## 9 Novel applications

Advances in FT-ICR MS capabilities and methodologies have significantly benefitted the research communities of geochemistry and atmospheric chemistry. For example, the three-dimensional distribution of DOM molecules can even be mapped and quantified by MS. Araújo et al. (2014) who applied a mass spectrometry imaging (MSI) technique to view the spatial and lateral distributions of soluble lignins in hand-cut sections of stems of Eucalyptus. As a result, the proportions of syringyl and guaiacyl monolignols were readily detected and relatively quantified (Fig. 8).

Molecular characterization of DOM is the first step to develop a complete understanding of environmental geochemistry processes and dynamics. Once the elemental composition is determined, the next step is to obtain its chemical and structural information. A thorough knowledge of DOM composition as well as its structure is crucial for the understanding of its role in the environment. Tandem mass spectrometry ( $MS^n$ ) and hydrogen/deuterium (H/D) exchange are two of the most important methods for structural analysis in the field of MS to offer further insights into the structural diversity.

For linear biomolecules such as proteins, their primary structures can be uniquely defined by the one-dimensional sequences, while for most DOM, structural elucidation requires multi-dimensional determinations, including functional groups, linkages, structural isomers and even stereochemical configurations. Tandem MS is a technique to break down selected ions (precursor ions) into small fragments (product ions) and in-depth structural information of the precursor ions can be proposed from the product ions.

Most commonly, collision induced dissociation (CID) is utilized as an ion activation method. In CID, precursor ions are isolated and subjected to collisions to induce fragmentations; the resulting product ions offer information on the ion's functional groups and structural units. The fragments can be further isolated and subjected to more collisions ( $MS^3$ ,  $MS^4$ ...) to acquire increased information (Sleno and Volmer 2004). CID with FT-ICR MS has gradually been applied to explore the dominant features of complex DOM mixtures, e.g.,





probe functional groups, CO<sub>2</sub> and H<sub>2</sub>O units originating from highly ionizable carboxylic groups (Witt et al. 2009; Zark et al. 2017; Zark and Dittmar 2018). Banoub et al. (2015) authored an informative tutorial article, which summarized the application of MS<sup>n</sup> for sequencing lignin and its degradation compounds. More recently, infrared multi-photon dissociation (IRMPD) has been introduced as a technique to cleave C-O and C-C bonds of a molecule's backbone chain, providing informative and complementary information to CID. Furthermore, IRMPD has been introduced to aquatic DOM to reveal structural information and reactivity pathways that are unavailable from CID (Brown et al. 2016; Kurek et al. 2020).

Although MS<sup>n</sup> plays an important role in protein sequencing, its application to DOM is currently still restricted to identify certain functional groups. This is caused by the complex features of DOM: e.g., ions species are too closely spaced to be individually isolated; isobaric interferences and precursor ions with the same nominal mass could be fragmented together in MS<sup>n</sup>, leading to ambiguous structural interpretation. Overall, efforts are required in the future to develop advanced approaches that are suitable for DOM structure analysis.

Normally, functional groups such as carboxyl, hydroxyl or amide contain labile hydrogens, which can be replaced by deuterium in solution or in the gas phase, and therefore, the number of hydrogens exchanged by deuterium in a DOM compound under such conditions might serve as an indicator to the number of the corresponding functional groups containing labile hydrogens. Kostyukevich et al. (2013) reported a simple in-source H/D exchange method based on FT-ICR MS for the determination of labile hydrogens in all constituents composing DOM. The experiment indicated that Suwannee River fulvic acid (SRFA) and lignosulfonate molecules contain 2-5 labile hydrogens, while the compounds of Siberian crude oil ionizing in negative ESI mode had one labile hydrogen. The three substances represent different degrees of transformation of DOM in nature, and H/D exchange FT-ICR MS offers new insights into the carbon cycle on the Earth.

The FT-ICR MS describes the DOM at the molecular level using chemical formulas. Unfortunately, the structural diversity of many compounds remains unknown. Moreover, the isomeric feature of DOM makes its structure analysis a challenging problem. In addition to MS<sup>n</sup> and H/D exchange, trapped ion mobility spectrometry (TIMS) offers an alternative to reveal the isomeric diversity and structure features of complex mixtures. For example, Gao et al. (2019) described the

SRNOM fractionated by multi-step SPE with different elution solvents, the fractions were both characterized by FT-ICR MS and IMS-MS, and showed the combination of the two instruments characterized the complex mixture into a new depth. Fernandez-Lima and co-workers pioneered the early works for the integration of TIMS with FT-ICR MS and reported its unique advantages for the analysis of complex mixtures (Fig. 9) (Benigni et al. 2017; Leyva et al. 2020).

It is also important to point out that TIMS-FT-ICR MS was able to report the lower abundance compositional trends compared to TIMS-TOF MS (Tose et al. 2018). Very recently, the Fernandez-Lima group further expanded the boundaries of TIMS-FT-ICR MS by performing elemental formula-based ion mobility and MS<sup>n</sup> analysis for DOM structural characterization (Leyva et al. 2020). In the reported work, ion species with a nominal mass of 393 were isolated including isobars and isomers, and five IMS bands were assigned to the heterogeneous profile of the precursor ion [C<sub>18</sub>H<sub>18</sub>O<sub>10</sub>-H]<sup>-</sup> and correlated with the data obtained by MS<sup>n</sup> (Fig. 9).

## 10 Current issues and future work

Mass spectrometry has become a well-developed and highly valuable tool in environmental and earth sciences. In early studies, FT-ICR MS was introduced to the geochemistry and atmospheric chemistry community to answer questions such as: What is the complex DOM mixture composed of? How to identify its sources and chemical transformations? How will it evolve and impact on the environment? During the last two decades, advances in the sample treatment, methodology development, data handling and instrument modification of FT-ICR MS have greatly contributed to a better understanding in the fields of environmental chemistry, organic geochemistry, and atmospheric chemistry.

Despite the fundamental principles, recent developments, and significant achievements of FT-ICR MS summarized here, the pertinent and immediate challenges remain. While ESI is the most viable method for DOM, various ionization methods are available to provide complementary information by preferentially accessing specific molecule types. Unfortunately, there are still fractions of DOM that cannot be fully detected. When the raw samples are solid or absorbed in situ, it is challenging to extract the representative DOM for the MS analysis. In addition, fragments and adducts resulting from the ionization process can further complicate the mass spectrum and lead to erroneous information. It is important to highlight that the signal intensity of compounds observed

in MS is influenced by a number of different factors, including sample preparation protocols, matrix effects, solvent pH values, molecules' ionization efficiencies, and instrument parameters. The contribution of these factors needs to be better considered, especially when comparing measurements from different batches and laboratories (Hawkes et al. 2020). More importantly, the observed compound intensity in a mass spectrum cannot be linked to its concentration straightforwardly. Although the use of spiked standards at known concentrations makes quantification of certain compounds of interest possible, quantification of a broader range of compounds presents a difficult task. Each analytical technique has its strengths and weaknesses, and the use of hyphenated techniques would address the remaining gaps and allow the most meaningful instrumental performance. Complementary analyses are increasingly used to better understand abiotic and biotic processes of DOM and to predict its geochemical impact.

When viewing the mass spectral data, it is essential to keep in mind that identical elemental formulas are not necessarily equal to identical molecules. Although FT-ICR MS is capable of doing compositional analysis of complex mixtures, illustration of the molecular structure remains a challenge. MS<sup>n</sup> approaches such as CID offer structural information; however, due to the complex nature of DOM samples, nominal mass resolution on the front-end quadrupole is not efficient enough to isolate a specific molecular ion from a narrow *m/z* window. Therefore, multiple ion species and their isomers would be mass fragmented simultaneously, leading to complicated MS<sup>n</sup> spectra, which are represented by several different precursor ions. Coupling with chromatography often allows separation of isomers and isobaric species, as well as efficient mass isolation of single molecular species. On the other hand, these hyphenated methodologies would generate much larger and even more complex datasets. Developing software for handling, visualization and comparison of multiple complex datasets is becoming a new challenge.

#### Authors' contributions

YQ, FW, GJ, CL and PF designed the review paper. YQ and PF wrote the paper. SL, JW, DH, HB, QF and DV helped to draft the manuscript and scientific discussions. QX, SS and MS edited the format and figures. All authors have given approval to the final version of the manuscript.

#### Funding

This research was funded by the National Natural Science Foundation of China (42007290 and 42130513). Dietrich A. Volmer acknowledges research support by the German Research Foundation (DVG VO 1355/4-3).

#### Availability of data and materials

The manuscript does not have any associated data.

## Declarations

#### Competing interests

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work done in this paper.

#### Author details

<sup>1</sup>Institute of Surface-Earth System Science, School of Earth System Science, Tianjin University, Tianjin, China. <sup>2</sup>Tianjin Key Laboratory of Earth Critical Zone Science and Sustainable Development in Bohai Rim, Tianjin University, Tianjin, China. <sup>3</sup>Haihe Laboratory of Sustainable Chemical Transformations, Tianjin, China. <sup>4</sup>State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China. <sup>5</sup>Department of Ocean Science and Hong Kong Branch of the Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), The Hong Kong University of Science and Technology, Hong Kong, China. <sup>6</sup>State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, China. <sup>7</sup>School of Environmental Studies, China University of Geosciences, Wuhan, China. <sup>8</sup>Department of Chemistry, Humboldt-Universität zu Berlin, Berlin, Germany. <sup>9</sup>State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing, China. <sup>10</sup>State Key Laboratory of Environmental Chemistry and Ecotoxicology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China.

Received: 29 March 2022 Accepted: 2 May 2022

Published online: 18 June 2022

## References

- Abe T, Maie N, Watanabe A (2005) Investigation of humic acid N with X-ray photoelectron spectroscopy: effect of acid hydrolysis and comparison with <sup>15</sup>N cross polarization/magic angle spinning nuclear magnetic resonance spectroscopy. *Org. Geochem.* 36(11):1490–1497
- Andreae M, Rosenfeld D (2008) Aerosol-cloud-precipitation interactions. Part 1. The nature and sources of cloud-active aerosols. *Earth-Sci. Rev.* 89(1–2):13–41
- Araújo P, MnS F, de Oliveira DN, Pereira L, Sawaya ACHF, Catharino RR, Mazzafera P (2014) Mass spectrometry imaging: an expeditious and powerful technique for fast in situ lignin assessment in Eucalyptus. *Anal. Chem.* 86(7):3415–3419
- Bahureksa W, Tfaily MM, Boiteau RM, Young RB, Logan MN, McKenna AM, Borch T (2021) Soil Organic Matter Characterization by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR MS): A Critical Review of Sample Preparation, Analysis, and Data Interpretation. *Environ. Sci. Technol.* 55(14):9637–9656
- Banoub J, Delmas GH Jr, Joly N, Mackenzie G, Cachet N, Benjelloun-Mlayah B, Delmas M (2015) A critique on the structural analysis of lignins and application of novel tandem mass spectrometric strategies to determine lignin sequencing. *J. Mass Spectrom.* 50(1):5–48
- Bao H, Niggemann J, Luo L, Dittmar T, Kao S-J (2018) Molecular composition and origin of water-soluble organic matter in marine aerosols in the Pacific off China. *Atmos. Environ.* 191:27–35
- Benigni P, Sandoval K, Thompson CJ, Ridgeway ME, Park MA, Gardinali P, Fernandez-Lima F (2017) Analysis of photoirradiated water accommodated fractions of crude oils using tandem TIMS and FT-ICR MS. *Environ. Sci. Technol.* 51(11):5978–5988
- Benner R, Newell S, Maccubbin A, Hodson RE (1984) Relative contributions of bacteria and fungi to rates of degradation of lignocellulosic detritus in salt-marsh sediments. *Appl. Environ. Microb.* 48(1):36–40
- Boesl U (2017) Time-of-flight mass spectrometry: introduction to the basics. *Mass Spectrom. Rev.* 36(1):86–109
- Brown TA, Jackson BA, Bythell BJ, Stenson AC (2016) Benefits of multidimensional fractionation for the study and characterization of natural organic matter. *J. Chromatogr. A* 1470:84–96
- Bruggemann M, Xu RS, Tilgner A, Kwong KC, Mutzel A, Poon HY, Otto T, Schaefer T, Poulain L, Chan MN, Herrmann H (2020) Organosulfates in

- Ambient Aerosol: State of Knowledge and Future Research Directions on Formation, Abundance, Fate, and Importance. *Environ. Sci. Technol.* 54(7):3767–3782
- Cao D, Huang HG, Hu M, Cui L, Geng FL, Rao ZY, Niu HY, Cai YQ, Kang YH (2015) Comprehensive characterization of natural organic matter by MALDI- and ESI-Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chim. Acta* 866:48–58
- Chen H, Munoz G, Duy SV, Zhang L, Yao Y, Zhao Z, Yi L, Liu M, Sun H, Liu J, Sauve S (2020) Occurrence and Distribution of Per- and Polyfluoroalkyl Substances in Tianjin, China: The Contribution of Emerging and Unknown Analogues. *Environ. Sci. Technol.* 54(22):14254–14264
- Chen S, Xie QR, Su SH, Wu LB, Zhong SJ, Zhang ZM, Ma C, Qi YL, Hu W, Deng JJ, Ren LJ, Zhu DQ, Guo QJ, Liu CQ, Jang KS, Fu PQ (2022) Source and formation process impact the chemodiversity of rainwater dissolved organic matter along the Yangtze River Basin in summer. *Water Res.* 211:118024
- Coble PG (2007) Marine optical biogeochemistry: the chemistry of ocean color. *Chem. Rev.* 107:402–418
- Comisarow MB, Marshall AG (1974) Fourier transform ion cyclotron resonance spectroscopy. *Chem. Phys. Lett.* 25(2):282–283
- Cory RM, McKnight DM (2005) Fluorescence spectroscopy reveals ubiquitous presence of oxidized and reduced quinones in dissolved organic matter. *Environ. Sci. Technol.* 39(21):8142–8149
- Da Silva MP, Kaesler JM, Reemtsma T, Lechtenfeld OJ (2020) Absorption mode spectral processing improves data quality of natural organic matter analysis by Fourier-transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.* 31(7):1615–1618
- D'Andrilli J, Chanton JP, Glaser PH, Cooper WT (2010a) Characterization of dissolved organic matter in northern peatland soil porewaters by ultrahigh resolution mass spectrometry. *Org. Geochem.* 41(8):791–799
- D'Andrilli J, Dittmar T, Koch BP, Purcell JM, Marshall AG, Cooper WT (2010b) Comprehensive characterization of marine dissolved organic matter by Fourier transform ion cyclotron resonance mass spectrometry with electrospray and atmospheric pressure photoionization. *Rapid Commun. Mass Spectrom.* 24(5):643–650
- D'Andrilli J, Cooper WT, Foreman CM, Marshall AG (2015) An ultrahigh-resolution mass spectrometry index to estimate natural organic matter lability. *Rapid Commun. Mass Spectrom.* 29(24):2385–2401
- Denisov E, Damoc E, Lange O, Makarov A (2012) Orbitrap mass spectrometry with resolving powers above 1,000,000. *Int. J. Mass Spectrom.* 325:80–85
- Dittmar T, Koch BP (2006) Thermogenic organic matter dissolved in the abyssal ocean. *Mar. Chem.* 102(3–4):208–217
- Dittmar T, Koch B, Hertkorn N, Kattner G (2008) A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr.-Meth.* 6(6):230–235
- Drosos M, Nebbioso A, Mazzei P, Vinci G, Spaccini R, Piccolo A (2017) A molecular zoom into soil Humeome by a direct sequential chemical fractionation of soil. *Sci. Total Environ.* 586:807–816
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1990) Electrospray ionization—principles and practice. *Mass Spectrom. Rev.* 9(1):37–70
- Fievre A, Solouki T, Marshall AG, Cooper WT (1997) High-resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids by laser desorption/ionization and electrospray ionization. *Energ. Fuel.* 11(3):554–560
- Fu PQ, Wu FC, Liu C-Q, Wei ZQ, Bai YC, Liao HQ (2006) Spectroscopic characterization and molecular weight distribution of dissolved organic matter in sediment porewaters from Lake Erhai, Southwest China. *Biogeochemistry* 81(2):179–189
- Fu PQ, Kawamura K, Chen J, Li J, Sun YL, Liu Y, Tachibana E, Aggarwal SG, Okuzawa K, Tanimoto H, Kanaya Y, Wang ZF (2012) Diurnal variations of organic molecular tracers and stable carbon isotopic composition in atmospheric aerosols over Mt. Tai in the North China Plain: an influence of biomass burning. *Atmos. Chem. Phys.* 12(18):8359–8375
- Fu Q-L, Fujii M, Riedel T (2020a) Development and comparison of formula assignment algorithms for ultrahigh-resolution mass spectra of natural organic matter. *Anal. Chim. Acta* 1125:247–257
- Fu Q-L, Fujii M, Kwon E (2020b) Development and Application of a High-Precision Algorithm for Nontarget Identification of Organohalogenes Based on Ultrahigh-Resolution Mass Spectrometry. *Anal. Chem.* 92(20):13989–13996
- Fu Q-L, Fujii M, Watanabe A, Kwon E (2022) Formula Assignment Algorithm for Deuterium-Labeled Ultrahigh-Resolution Mass Spectrometry: Implications of the Formation Mechanism of Halogenated Disinfection Byproducts. *Anal. Chem.* 94(3):1717–1725
- Gao Y, Wang W, He C, Fang Z, Zhang Y, Shi Q (2019) Fractionation and molecular characterization of natural organic matter (NOM) by solid-phase extraction followed by FT-ICR MS and ion mobility MS. *Anal. Bioanal. Chem.* 411(24):6343–6352
- Ge J, Qi Y, Li C, Ma J, Yi Y, Hu Q, Mostofa KMG, Volmer DA, Li SL (2022) Fluorescence and molecular signatures of dissolved organic matter to monitor and assess its multiple sources from a polluted river in the farming-pastoral ecotone of northern China. *Sci. Total Environ.* <https://doi.org/10.1016/j.scitotenv.2022.154575>
- Goñi MA, Hedges JI (1992) Lignin dimers: Structures, distribution, and potential geochemical applications. *Geochim. Cosmochim. Acta* 56(11):4025–4043
- Han HX, Feng YJ, Chen J, Xie QR, Chen S, Sheng M, Zhong SJ, Wei W, Su SH, Fu PQ (2022) Acidification impacts on the molecular composition of dissolved organic matter revealed by FT-ICR MS. *Sci. Total Environ.* 805:150284
- Hawkes JA, Patriarca C, Sjöberg PJ, Tranvik LJ, Bergquist J (2018) Extreme isomeric complexity of dissolved organic matter found across aquatic environments. *Limnol. Oceanogr. Lett.* 3(2):21–30
- Hawkes JA, Sjöberg PJ, Bergquist J, Tranvik L (2019) Complexity of dissolved organic matter in the molecular size dimension: insights from coupled size exclusion chromatography electrospray ionisation mass spectrometry. *Faraday Discuss.* 218:52–71
- Hawkes JA, d'Andrilli J, Agar JN, Barrow MP, Berg SM, Catalán N, Chen H, Chu RK, Cole RB, Dittmar T (2020) An international laboratory comparison of dissolved organic matter composition by high resolution mass spectrometry: Are we getting the same answer? *Limnol. Oceanogr.-Meth.* 18(6):235–258
- He D, He C, Li P, Zhang X, Shi Q, Sun Y (2019) Optical and Molecular Signatures of Dissolved Organic Matter Reflect Anthropogenic Influence in a Coastal River, Northeast China. *J. Environ. Qual.* 48(3):603–613
- He D, Wang K, Pang Y, He C, Li P, Li Y, Xiao S, Shi Q, Sun Y (2020) Hydrological management constraints on the chemistry of dissolved organic matter in the Three Gorges Reservoir. *Water Res.* 187:116413
- Hedges JI, Ertel JR, Leopold EB (1982) Lignin geochemistry of a Late Quaternary sediment core from Lake Washington. *Geochim. Cosmochim. Acta* 46(10):1869–1877
- Hedges JI, Eglinton G, Hatcher PG, Kirchman DL, Arnosti C, Derenne S, Evershed RP, Kögel-Knabner I, de Leeuw JW, Littke R (2000) The molecularly-uncharacterized component of nonliving organic matter in natural environments. *Org. Geochem.* 31(10):945–958
- Hertkorn N, Benner R, Frommberger M, Schmitt-Kopplin P, Witt M, Kaiser K, Ketttrup A, Hedges JI (2006) Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Acta* 70(12):2990–3010
- Hsu CS, Hendrickson CL, Rodgers RP, McKenna AM, Marshall AG (2011) Petroleomics: advanced molecular probe for petroleum heavy ends. *J. Mass Spectrom.* 46(4):337–343
- Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R (2005) The Orbitrap: a new mass spectrometer. *J. Mass Spectrom.* 40(4):430–443
- Hu W, Liu DD, Su SH, Ren LJ, Ren H, Wei LF, Yue SY, Xie QR, Zhang ZM, Wang ZH, Yang N, Wu LB, Deng JJ, Qi YL, Fu PQ (2021) Photochemical degradation of organic matter in the atmosphere. *Adv. Sustain. Syst.* 5(11):2100027
- Hur M, Ware RL, Park J, McKenna AM, Rodgers RP, Nikolau BJ, Wurtele ES, Marshall AG (2018) Statistically significant differences in composition of petroleum crude oils revealed by volcano plots generated from ultrahigh resolution Fourier transform ion cyclotron resonance mass spectra. *Energ. Fuel.* 32(2):1206–1212
- Kew W, Blackburn JW, Clarke DJ, Uhrin D (2017) Interactive van Krevelen diagrams—Advanced visualisation of mass spectrometry data of complex mixtures. *Rapid Commun. Mass Spectrom.* 31(7):658
- Kim S, Kramer RW, Hatcher PG (2003) Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. *Anal. Chem.* 75(20):5336–5344

- Kim D, Lee J, Kim B, Kim S (2018) Optimization and application of paper-based spray ionization mass spectrometry for analysis of natural organic matter. *Anal. Chem.* 90(20):12027–12034
- Kim D, Kim S, Son S, Jung M-J, Kim S (2019) Application of online liquid chromatography 7 T FT-ICR mass spectrometer equipped with quadrupole detection for analysis of natural organic matter. *Anal. Chem.* 91(12):7690–7697
- Koch BP, Dittmar T (2006) From mass to structure: An aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* 20(5):926–932
- Koch BP, Witt M, Engbrodt R, Dittmar T, Kattner G (2005) Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochim. Cosmochim. Acta* 69(13):3299–3308
- Koch BP, Ludwiczowski 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. *Mar. Chem.* 111(3–4):233–241
- Kostyukovich Y, Kononikhin A, Popov I, Kharybin O, Perminova I, Konstantinov A, Nikolaev E (2013) Enumeration of labile hydrogens in natural organic matter by use of hydrogen/deuterium exchange Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 85(22):11007–11013
- Kundu S, Fisseha R, Putman AL, Rahn TA, Mazzoleni LR (2012) High molecular weight SOA formation during limonene ozonolysis: insights from ultrahigh-resolution FT-ICR mass spectrometry characterization. *Atmos. Chem. Phys.* 12(12):5523–5536
- Kurek MR, Poulin BA, McKenna AM, Spencer RG (2020) Deciphering Dissolved Organic Matter: Ionization, Dopant, and Fragmentation Insights via Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry. *Environ. Sci. Technol.* 54(24):16249–16259
- Laskin A, Laskin J, Nizkorodov SA (2015) Chemistry of atmospheric brown carbon. *Chem. Rev.* 115(10):4335–4382
- LeClair JP, Collett JL, Mazzoleni LR (2012) Fragmentation Analysis of Water-Soluble Atmospheric Organic Matter Using Ultrahigh-Resolution FT-ICR Mass Spectrometry. *Environ. Sci. Technol.* 46(8):4312–4322
- Leefmann T, Frickenhaus S, Koch BP (2019) UltraMassExplorer: a browser-based application for the evaluation of high-resolution mass spectrometric data. *Rapid Commun. Mass Spectrom.* 33(2):193–202
- Leyva D, Tose LV, Porter J, Wolff J, Jaffé R, Fernandez-Lima F (2019) Understanding the structural complexity of dissolved organic matter: isomeric diversity. *Faraday Discuss.* 218:431–440
- Leyva D, Jaffe R, Fernandez-Lima F (2020) Structural characterization of dissolved organic matter at the chemical formula level using TIMS-FT-ICR MS/MS. *Anal. Chem.* 92(17):11960–11966
- Li P, Hur J (2017) Utilization of UV-Vis spectroscopy and related data analyses for dissolved organic matter (DOM) studies: A review. *Crit. Rev. Environ. Sci. Technol.* 47(3):131–154
- Li X-M, Sun G-X, Chen S-C, Fang Z, Yuan H-Y, Shi Q, Zhu Y-G (2018) Molecular chemodiversity of dissolved organic matter in paddy soils. *Environ. Sci. Technol.* 52(3):963–971
- Li JJ, Zhang Q, Wang GH, Li J, Wu C, Liu L, Wang JY, Jiang WQ, Li LJ, Ho KF, Cao JJ (2020) Optical properties and molecular compositions of water-soluble and water-insoluble brown carbon (BrC) aerosols in northwest China. *Atmos. Chem. Phys.* 20(8):4889–4904
- Lin P, Rincon AG, Kalberer M, Yu JZ (2012) Elemental composition of HULIS in the Pearl River Delta Region, China: results inferred from positive and negative electrospray high resolution mass spectrometric data. *Environ. Sci. Technol.* 46(14):7454–7462
- Louchouart P, Amon RM, Duan S, Pondell C, Seward SM, White N (2010) Analysis of lignin-derived phenols in standard reference materials and ocean dissolved organic matter by gas chromatography/tandem mass spectrometry. *Mar. Chem.* 118(1–2):85–97
- Lozano DCP, Gavard R, Arenas-Diaz JP, Thomas MJ, Stranz DD, Mejía-Ospino E, Guzman A, Spencer SE, Rossell D, Barrow MP (2019) Pushing the analytical limits: new insights into complex mixtures using mass spectra segments of constant ultrahigh resolving power. *Chem. Sci.* 10(29):6966–6978
- Lozano DCP, Thomas MJ, Jones HE, Barrow MP (2020) Petroleomics: Tools, Challenges, and Developments. In: Bohn PW, Pemberton JE (eds) *Annual Review of Analytical Chemistry*, Vol 13, *Ann. Rev. Anal. Chem.*, vol 13, pp 405–430
- Mann BF, Chen H, Herndon EM, Chu RK, Tolic N, Portier EF, Roy Chowdhury T, Robinson EW, Callister SJ, Wullschlegel SD (2015) Indexing permafrost soil organic matter degradation using high-resolution mass spectrometry. *PLoS One* 10(6):e0130557
- Merder J, Freund JA, Feudel U, Niggemann J, Singer G, Dittmar T (2019) Improved mass accuracy and isotope confirmation through alignment of ultrahigh-resolution mass spectra of complex natural mixtures. *Anal. Chem.* 92(3):2558–2565
- Merder J, Freund JA, Feudel U, Hansen CT, Hawkes JA, Jacob B, Klapproth K, Niggemann J, Noriega-Ortega BE, Osterholz H (2020) ICBM-OCEAN: processing ultrahigh-resolution mass spectrometry data of complex molecular mixtures. *Anal. Chem.* 92(10):6832–6838
- Minor EC, Swenson MM, Mattson BM, Oyler AR (2014) Structural characterization of dissolved organic matter: a review of current techniques for isolation and analysis. *Environ. Sci. Proc. Imp.* 16(9):2064–2079
- Mopper K, Stubbins A, Ritchie JD, Bialk HM, Hatcher PG (2007) Advanced instrumental approaches for characterization of marine dissolved organic matter: extraction techniques, mass spectrometry, and nuclear magnetic resonance spectroscopy. *Chem. Rev.* 107(2):419–442
- Nebbioso A, Piccolo A (2013) Molecular characterization of dissolved organic matter (DOM): a critical review. *Anal. Bioanal. Chem.* 405(1):109–124
- Nelson NB, Siegel DA (2013) The global distribution and dynamics of chromophoric dissolved organic matter. *Ann. Rev. Mar. Sci.* 5:447–476
- Novak J, Bertsch P, Mills G (1992) Carbon-13 Nuclear Magnetic Resonance Spectra of Soil Water-Soluble Organic Carbon. *J. Environ. Qual.* 21(4):537–539
- Opsahl S, Benner R (1997) Distribution and cycling of terrigenous dissolved organic matter in the ocean. *Nature* 386(6624):480–482
- Pang Y, Wang K, Sun Y, Zhou Y, Yang S, Li Y, He C, Shi Q, He D (2021) Linking the unique molecular complexity of dissolved organic matter to flood period in the Yangtze River mainstream. *Sci Total Environ.* 764:142803
- Podgorski DC, McKenna AM, Rodgers RP, Marshall AG, Cooper WT (2012) Selective ionization of dissolved organic nitrogen by positive ion atmospheric pressure photoionization coupled with Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 84(11):5085–5090
- Qi YL, Barrow MP, Li H, Meier JE, Van Orden SL, Thompson CJ, O'Connor PB (2012) Absorption-mode: the next generation of Fourier transform mass spectra. *Anal. Chem.* 84(6):2923–2929
- Qi YL, Fu PQ, Li S, Ma C, Liu C, Volmer DA (2020) Assessment of molecular diversity of lignin products by various ionization techniques and high-resolution mass spectrometry. *Sci. Total Environ.* 713:136573
- Qi YL, Fu PQ, Volmer DA (2022) Analysis of natural organic matter via Fourier transform ion cyclotron resonance mass spectrometry: an overview of recent non-petroleum applications. *Mass Spectrom. Rev.* <https://doi.org/10.1002/mas.21634>
- Qi YL, Hempelmann R, Volmer DA (2016) Two-dimensional mass defect matrix plots for mapping genealogical links in mixtures of lignin depolymerization products. *Anal. Bioanal. Chem.* 408(18):4835–4843
- Qi YL, Ma C, Chen S, Ge JF, Hu QZA, Li SL, Volmer DA, Fu PQ (2021) Online Liquid Chromatography and FT-ICR MS Enable Advanced Separation and Profiling of Organosulfates in Dissolved Organic Matter. *ACS ES&T Water* 1(8):1975–1982
- Qi YL, O'Connor PB (2014) Data processing in Fourier transform ion cyclotron resonance mass spectrometry. *Mass Spectrom. Rev.* 33(5):333–352
- Qi YL, Volmer DA (2019) Chemical diversity of lignin degradation products revealed by matrix-optimized MALDI mass spectrometry. *Anal. Bioanal. Chem.* 411(23):6031–6037
- Riedel T, Dittmar T (2014) A method detection limit for the analysis of natural organic matter via Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 86(16):8376–8382
- Rivas-Ubach A, Liu Y, Bianchi TS, Tolic N, Jansson C, Pasa-Tolic L (2018) Moving beyond the van Krevelen diagram: A new stoichiometric approach for compound classification in organisms. *Anal. Chem.* 90(10):6152–6160
- Roebuck JA Jr, Seidel M, Dittmar T, Jaffé R (2019) Controls of land use and the river continuum concept on dissolved organic matter composition in an anthropogenically disturbed subtropical watershed. *Environ. Sci. Technol.* 54(1):195–206
- Roth V-N, Lange M, Simon C, Hertkorn N, Bucher S, Goodall T, Griffiths RI, Mel-lado-Vázquez PG, Mommer L, Oram NJ, Weigelt A, Dittmar T, Gleixner G



- (2019) Persistence of dissolved organic matter explained by molecular changes during its passage through soil. *Nature Geosci.* 12(9):755–761
- Sandron S, Rojas A, Wilson R, Davies NW, Haddad PR, Shellie RA, Nesterenko PN, Kelleher BP, Paull B (2015) Chromatographic methods for the isolation, separation and characterisation of dissolved organic matter. *Environ. Sci. Proc-Imp.* 17(9):1531–1567
- Schmidt MW, Torn MS, Abiven S, Dittmar T, Guggenberger G, Janssens IA, Kleber M, Kögel-Knabner I, Lehmann J, Manning DA (2011) Persistence of soil organic matter as an ecosystem property. *Nature* 478(7367):49–56
- Schum SK, Brown LE, Mazzoleni LR (2020) MFAAssignR: Molecular formula assignment software for ultrahigh resolution mass spectrometry analysis of environmental complex mixtures. *Environ. Res.* 191:110114
- Schuur EA, Abbott B, Bowden W, Brovkin V, Camill P, Canadell J, Chanton J, Chapin F, Christensen T, Ciais P (2013) Expert assessment of vulnerability of permafrost carbon to climate change. *Climatic Change* 119(2):359–374
- Simjouw J-P, Minor EC, Mopper K (2005) Isolation and characterization of estuarine dissolved organic matter: comparison of ultrafiltration and C18 solid-phase extraction techniques. *Mar. Chem.* 96(3–4):219–235
- Sleighter RL, Hatcher PG (2007) The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *J. Mass Spectrom.* 42(5):559–574
- Sleighter RL, Hatcher PG (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. *Mar. Chem.* 110(3–4):140–152
- Sleno L, Volmer DA (2004) Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* 39(10):1091–1112
- Smith DF, Podgorski DC, Rodgers RP, Blakney GT, Hendrickson CL (2018) 21 tesla FT-ICR mass spectrometer for ultrahigh-resolution analysis of complex organic mixtures. *Anal. Chem.* 90(3):2041–2047
- Solihat NN, Acter T, Kim D, Plante AF, Kim S (2018) Analyzing solid-phase natural organic matter using laser desorption ionization ultrahigh resolution mass spectrometry. *Anal. Chem.* 91(1):951–957
- Stedmon CA, Bro R (2008) Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. *Limnol. Oceanogr-Meth.* 6(11):572–579
- Stenson AC, Marshall AG, Cooper WT (2003) Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Anal. Chem.* 75(6):1275–1284
- Stubbins A, Lapierre J-F, Berggren M, Prairie YT, Dittmar T, del Giorgio PA (2014) What's in an EEM? Molecular signatures associated with dissolved organic fluorescence in boreal Canada. *Environ. Sci. Technol.* 48(18):10598–10606
- Su SH, Xie QR, Lang YC, Cao D, Xu YS, Chen J, Chen S, Hu W, Qi YL, Pan XL, Sun YL, Wang ZF, Liu CQ, Jiang GB, Fu PQ (2021) High Molecular Diversity of Organic Nitrogen in Urban Snow in North China. *Environ. Sci. Technol.* 55(8):4344–4356
- Sullivan A, Weber R, Clements A, Turner J, Bae M, Schauer J (2004) A method for on-line measurement of water-soluble organic carbon in ambient aerosol particles: Results from an urban site. *Geophys. Res. Lett.* 31(13):L13105. <https://doi.org/10.1029/2004GL019681>
- Tang J, Li J, Su T, Han Y, Mo YZ, Jiang HX, Cui M, Jiang B, Chen YJ, Tang JH, Song JZ, Peng PA, Zhang G (2020) Molecular compositions and optical properties of dissolved brown carbon in biomass burning, coal combustion, and vehicle emission aerosols illuminated by excitation-emission matrix spectroscopy and Fourier transform ion cyclotron resonance mass spectrometry analysis. *Atmos. Chem. Phys.* 20(4):2513–2532
- Tfaily MM, Hamdan R, Corbett JE, Chanton JP, Glaser PH, Cooper WT (2013) Investigating dissolved organic matter decomposition in northern peatlands using complementary analytical techniques. *Geochim. Cosmochim. Acta* 112:116–129
- Tfaily MM, Chu RK, Tolić N, Roscioli KM, Anderton CR, Paša-Tolić L, Robinson EW, Hess NJ (2015) Advanced solvent based methods for molecular characterization of soil organic matter by high-resolution mass spectrometry. *Anal. Chem.* 87(10):5206–5215
- Tolić N, Liu Y, Liyu A, Shen Y, Tfaily MM, Kujawinski EB, Longnecker K, Kuo L-J, Robinson EW, Paša-Tolić L, Hess NJ (2017) Formularity: Software for Automated Formula Assignment of Natural and Other Organic Matter from Ultrahigh-Resolution Mass Spectra. *Anal. Chem.* 89(23):12659–12665
- Tose LV, Benigni P, Leyva D, Sundberg A, Ramírez CE, Ridgeway ME, Park MA, Romão W, Jaffé R, Fernandez-Lima F (2018) Coupling trapped ion mobility spectrometry to mass spectrometry: trapped ion mobility spectrometry-time-of-flight mass spectrometry versus trapped ion mobility spectrometry-Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.* 32(15):1287–1295
- Tranvik LJ, Downing JA, Cotner JB, Loiselle SA, Striegl RG, Ballatore TJ, Dillon P, Finlay K, Fortino K, Knoll LB (2009) Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol. Oceanogr.* 54(6):2298–2314
- Tremblay LB, Dittmar T, Marshall AG, Cooper WJ, Cooper WT (2007) Molecular characterization of dissolved organic matter in a North Brazilian mangrove porewater and mangrove-fringed estuaries by ultrahigh resolution Fourier Transform-Ion Cyclotron Resonance mass spectrometry and excitation/emission spectroscopy. *Mar. Chem.* 105(1–2):15–29
- Vonk JE, Mann PJ, Davydov S, Davydova A, Spencer RG, Schade J, Sobczak WV, Zimov N, Zimov S, Bulygina E (2013) High biolability of ancient permafrost carbon upon thaw. *Geophys. Res. Lett.* 40(11):2689–2693
- Wagner S, Riedel T, Niggemann J, Vähätalo AV, Dittmar T, Jaffé R (2015) Linking the molecular signature of heteroatomic dissolved organic matter to watershed characteristics in world rivers. *Environ. Sci. Technol.* 49(23):13798–13806
- Wang J-J, Liu Y, Bowden RD, Lajtha K, Simpson AJ, Huang W-L, Simpson MJ (2019a) Long-Term Nitrogen Addition Alters the Composition of Soil-Derived Dissolved Organic Matter. *ACS Earth and Space Chem.* 4(2):189–201
- Wang K, Pang Y, He C, Li P, Xiao S, Sun Y, Pan Q, Zhang Y, Shi Q, He D (2019b) Optical and molecular signatures of dissolved organic matter in Xiangxi Bay and mainstream of Three Gorges Reservoir, China: Spatial variations and environmental implications. *Sci. Total Environ.* 657:1274–1284
- Wang K, Li P, He C, Shi Q, He D (2021a) Hydrologic heterogeneity induced variability of dissolved organic matter chemistry among tributaries of the Three Gorges Reservoir. *Water Res.* 201:117358
- Wang K, Pang Y, Li Y, He C, Shi Q, Wang Y, He D (2021b) Characterizing Dissolved Organic Matter Across a Riparian Soil–Water Interface: Preliminary Insights from a Molecular Level Perspective. *ACS Earth and Space Chem.* 5(5):1102–1113
- Wang Y, Zhang Q, Majidzadeh H, He C, Shi Q, Kong S, Yang Z, Wang J (2021c) Depletion of Soil Water-Extractable Organic Matter With Long-Term Coverage by Impervious Surfaces. *Front. Environ. Sci.* 9:290
- Witt M, Fuchser J, Koch BP (2009) Fragmentation studies of fulvic acids using collision induced dissociation Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 81(7):2688–2694
- Wu FC, Tanoue E (2001) Isolation and partial characterization of dissolved copper-complexing ligands in streamwaters. *Environ. Sci. Technol.* 35(18):3646–3652
- Wu LB, Yue SY, Shi ZB, Hu W, Chen J, Ren H, Deng JJ, Ren LJ, Fang YT, Yan H, Li WJ, Harrison RM, Fu PQ (2021) Source forensics of inorganic and organic nitrogen using  $\delta^{15}\text{N}$  for tropospheric aerosols over Mt. Tai. *npj Clim. Atmos. Sci.* 4(1):1–8
- Xie QR, Li Y, Yue SY, Su SH, Cao D, Xu YS, Chen J, Tong HJ, Su H, Cheng YF, Zhao WY, Hu W, Wang Z, Yang T, Pan XL, Sun YL, Wang ZF, Liu CQ, Kawamura K, Jiang GB, Shiraiwa M, Fu PQ (2020) Increase of High Molecular Weight Organosulfate With Intensifying Urban Air Pollution in the Megacity Beijing. *J. Geophys. Res-Atmos.* 125(10):10.1029/2019jd032200
- Xu C, Gao LR, Zheng MH, Qiao L, Wang KR, Huang D, Wang S (2021) Nontarget Screening of Polycyclic Aromatic Compounds in Atmospheric Particulate Matter Using Ultrahigh Resolution Mass Spectrometry and Comprehensive Two-Dimensional Gas Chromatography. *Environ. Sci. Technol.* 55(1):109–119
- Yamashita Y, Jaffé R (2008) Characterizing the interactions between trace metals and dissolved organic matter using excitation-emission matrix and parallel factor analysis. *Environ. Sci. Technol.* 42(19):7374–7379
- Ye Q, Wang Y-H, Zhang Z-T, Huang W-L, Li L-P, Li J, Liu J, Zheng Y, Mo J-M, Zhang W, Wang J-J (2020) Dissolved organic matter characteristics in soils of tropical legume and non-legume tree plantations. *Soil Biol. Biochem.* 148:107880
- Yue SY, Zhu JL, Chen S, Xie QR, Li W, Li LJ, Ren H, Su SH, Li P, Ma H, Fan YB, Cheng BR, Wu LB, Deng JJ, Hu W, Ren LJ, Wei LF, Zhao WY, Tian Y, Pan

- XL, Sun YL, Wang ZF, Wu FC, Liu C-Q, Su H, Penner JE, Pöschl U, Andreae MO, Cheng YF, Fu PQ (2022) Brown carbon from biomass burning imposes strong circum-Arctic warming. *One Earth* 5(3):293–304
- Zark M, Dittmar T (2018) Universal molecular structures in natural dissolved organic matter. *Nat. Commun.* 9(1):1–8
- Zark M, Christoffers J, Dittmar T (2017) Molecular properties of deep-sea dissolved organic matter are predictable by the central limit theorem: Evidence from tandem FT-ICR-MS. *Mar. Chem.* 191:9–15
- Zhang HF, Zhang YH, Shi Q, Ren SY, Yu JW, Ji F, Luo WB, Yang M (2012) Characterization of low molecular weight dissolved natural organic matter along the treatment trail of a waterworks using Fourier transform ion cyclotron resonance mass spectrometry. *Water Res.* 46(16):5197–5204
- Zhang X, Chen Z, Xiaoyu H, Kang J, Zhao S, Peng Y, Deng F, Shen J, Chu W (2020a) Application of Fourier transform ion cyclotron resonance mass spectrometry in deciphering molecular composition of soil organic matter: A review. *Sci. Total Environ.* 756:144140
- Zhang X, Han J, Zhang X, Shen J, Chen Z, Chu W, Kang J, Zhao S, Zhou Y (2020b) Application of Fourier transform ion cyclotron resonance mass spectrometry to characterize natural organic matter. *Chemosphere* 260:127458
- Zhang R, Qi Y, Ma C, Ge J, Hu Q, Yue F-J, Li S-L, Volmer DA (2021) Characterization of Lignin Compounds at the Molecular Level: Mass Spectrometry Analysis and Raw Data Processing. *Molecules* 26(1):178
- Zhao Y, Hallar AG, Mazzoleni LR (2013) Atmospheric organic matter in clouds: exact masses and molecular formula identification using ultrahigh-resolution FT-ICR mass spectrometry. *Atmos. Chem. Phys.* 13(24):12343–12362
- Zhou YQ, Davidson TA, Yao XL, Zhang YL, Jeppesen E, de Souza JG, Wu HW, Shi K, Qin BQ (2018) How autochthonous dissolved organic matter responds to eutrophication and climate warming: Evidence from a cross-continental data analysis and experiments. *Earth-Sci. Rev.* 185:928–937
- Zhou YQ, Zhou L, Zhang YL, Zhu GW, Qin BQ, Jang KS, Spencer RGM, Kothawala DN, Jeppesen E, Brookes JD, Wu FC (2022) Unraveling the Role of Anthropogenic and Natural Drivers in Shaping the Molecular Composition and Biolability of Dissolved Organic Matter in Non-pristine Lakes. *Environ. Sci. Technol.* 56(7):4655–4664

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.