Environmental Science
Processes & Impacts

ROYAL SOCIETY
OF CHEMISTRY

rsc.li/process-impacts

Environmental Impact

In aquatic systems, dissolved organic matter (DOM) acts as a sunscreen, a food source, a trace metal chelator, and a photosensitizer. DOM's composition, which varies across environments, affects its ability to play such roles, but is difficult to ascertain because of challenges in isolating and analyzing it. In many studies solid-phase extraction has been used to isolate DOM, and it is ideally suited for coupling with mass spectrometry. This study assesses the performance of 2 solid-phase resins on the extraction of DOM from Lake Superior, the world's largest lake by area. Both sample recovery and the molecular-level fractionation of the isolated DOC were investigated, thus providing critical information for the comparison of different DOM-related environmental studies in natural systems.

Dissolved organic matter in Lake Superior: insights into the effects of extraction methods on chemical composition

Hongyu Li [a,]*, Elizabeth C. Minor [b]

Abstract Dissolved organic matter (DOM) in aquatic systems plays many biogeochemical roles, acting as a sunscreen, a food source, a trace metal chelator, and a photosensitizer. The efficiency of DOM in these roles is, in part, a function of its composition, which is difficult to determine due to its heterogeneity and the difficulty in isolating representative portions for subsequent molecular level analyses. In this study, the performance of two major types of solid phase extraction (SPE) resins (C18 vs polymeric SDVB) in disk format (C18 disk vs SDB-XC disk) was studied using DOM from Lake Superior, the earth's largest lake by area. The performance of the two SPE disks and their influences on the molecular chemical composition of the extracted retentates were studied with Uv-vis spectrometry and negative-ion electrospray Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS). We found that SDB-XC disks outperformed C18 disks in the isolation of DOC in terms of both higher recovery and less fractionation relative to the initial DOM composition. Extracts of the same samples obtained with the different resins shared 70% of molecular formulae. Compounds exclusive to the SDB-XC extractions exhibited similar compound distributions to those of the shared formulae but were somewhat contain more N, P or S and more aromatic. The C18 exclusive compounds had somewhat higher H/C ratios and contained a large proportion of compounds with oxygen and nitrogen (CHON). Cluster analysis and principle component analysis confirmed that sample location was the main driver of the composition of extracted samples but showed some fractionation of the samples based upon the type of resin.

1. Introduction

Dissolved organic matter (DOM) is ubiquitous in aquatic systems, and it constitutes one of the largest dynamic reservoirs of organic carbon on Earth.[1] Its variable molar masses and chemical structures help to determine its roles in the natural environment, including acting as a food and nutrient source to aquatic organisms, as a photosensitizer for anthropogenic compounds, and as a chelator of trace metals. Identifying the molecular composition of DOM, especially the molecularly uncharacterized fraction, is fundamental to the understanding of the sources, reactivity and cycling of DOM as well as global C.[2] However, the very low concentration of organic matter (~1 to 2 mg C/l) in the oceans and oligotrophic lakes [3,4,5] challenges most analytical techniques, especially for marine water which contains large amounts of inorganic salts.

A major focus of aquatic DOM research has been in the isolation and desalting of sufficient amounts of representative material that would provide more comprehensive structural information using analytical methods such as nuclear magnetic resonance (NMR), infrared spectroscopy and mass spectrometry. Such isolation approaches fall into the three main categories: solid phase extraction (SPE), ultrafiltration, and reverse osmosis and reverse osmosis/electrodialysis (RO/ED). The SPE approach for the isolation of DOM using one or a combination of XAD resins was established in the late 1970s and has been applied to many studies as the classic extraction approach.[6] Very different recoveries were obtained for different aquatic systems, e.g., a total recovery of 16-21% for sea water [7] and 50 to 80% recoveries for "colored water" with a high contribution from humic substances.[8] This is possibly due to a selective isolation of fractions of DOM with these resins.[9] Recently, C18 resin, which is a silica based octadecyl carbon sorbents, and polymeric styrene-divinylbenzene (SDVB) resin have been widely applied [10,11,12], usually to acidified samples. A comparison of the two types of resins in cartridge format to acidified water samples (primarily marine but including a groundwater sample) showed that C18 resins (with 25 to 40% recovery) were the most efficient silica-based sorbents, while resins based upon styrene divinyl benzene polymers (e.g., PPL, Oasis HLB, LiChrolut EN) outperformed C18 resins, with PPL resin producing the highest DOM recoveries (up to ~62% recovery).[9,13]

SPE resins have been found to selectively isolate certain fractions of compounds in DOM as illustrated in recent work examining SPE fractionation with subsequent characterization by high-resolution mass spectrometry or NMR.[12,14] These studies show that C18 resin discriminates against more oxygenated compounds (tannins?) as well as aliphatic amines and amides, fractionates the carboxylic acid pool [14], and, compared with other tested resins, preferentially retains higher H/C compounds [12]. A comparison of the retention of C18 vs PPL resins showed that C18 resin retained more saturated aliphatics and PPL retained more material falling within van Krevelen ranges seen for protein and CRAM, while NMR spectra of the same samples isolated by the two techniques looked very similar.[12]

Despite SPE drawbacks, such as fractionation, and the possibility of chemical changes within the extracts due to pH alterations of DOM structure, SPE retains a significant role in DOM concentration and isolation studies due to its simple setup and wide applicability in different working conditions, especially in remote field settings.

Ultrafiltration, which isolates DOM fractions based on molecular size (primarily with a 1-kDa membrane), is another widely recognized method for the extraction of DOM (ultrafiltered DOM, or UDOM), and the one most favored by isotope geochemists. It isolates up to 30% of marine DOM and larger percentages of the initial DOC in colored waters, such as most lakes and rivers which have larger terrestrial impact, and higher average molecular weight DOM.[15,16] The combination of reverse osmosis and electrodialysis allows isolation of much larger proportions of aquatic DOM as compared to other techniques, with DOC recovery ranging 60% to 95%.[17,18] However, the RO/ED technique requires access to expensive equipment that can be prohibitive in many cases especially in the field, and due to large membrane surface areas, may contribute significantly to organic carbon blanks.

Advanced spectroscopic techniques have been employed to identify the composition of DOM in both raw water and concentrates. Measurements of DOC and colored DOM (via UV-spectrometry) provide average evaluation of the bulk properties of each DOM sample. FT-ICR-MS, which has a high degree of mass accuracy and precision, can provide molecular information on complex natural OM and has been yielding new insights into DOM variations in aquatic systems.[19,20,21,22] FT-ICR-MS can detect ions over a wide range of mass to charge ratios (m/z) at high accuracy, coupled with electrospray ionization (ESI), which ionizes both acidic and basic functional groups in DOM (depending upon ESI settings and sample solution pH) at atmospheric pressure, a wide range of molecules in DOM can be detected.[23, 24] A typical DOM mass spectrum results in a large matrix of thousands of molecular formulae and their signal intensities. Thus comparing multiple samples to understand environmental variations in DOM structure requires appropriate data visualization and data mining approaches.[23,24,25]

In this study, the performance of C18 and SDB-XC disks was compared using coastal and off-shore Lake Superior water. Dissolved organic carbon (DOC) concentrations, stable ($\delta^{13}C$) and radiocarbon (^{14}C) signatures and UV-Visible analyses of bulk DOM samples were used to assess the nature of water samples from different sites and depths. DOC and UV-Visible analyses of raw water, filtrates and retentates (DOM isolates) were also used to evaluate disk performance, including both the quantity and quality of isolated DOM. The molecular level composition of the extracted DOM obtained using the different disk resins (C18 vs SDB-XC) from one offshore site (WM) and one nearshore site (BR) were investigated using ESI FT-ICR MS. We chose to test C18 resin (C18 disk) and SDVB resin (SDB-XC disk) in disk format as this format is marketed as more appropriate for high pressure and large volume extractions than cartridges; disks also require less solvent use. The C18 disk approach has been applied for decades, while SDB-XC disk use is relatively new. As SDB-XC uses an SDVB resin construction it should perform similarly to PPL cartridges, which have been reported to produce higher DOC recovery.[9] This comparison may provide insight into the impact of two major SPE resin types to

DOM extraction and assist comparison of current and previous studies sampled with different resins (C18 vs SDVB).

2. Experimental

2.1. Sites and Sampling

Lake Superior is the Earth's largest lake by surface area (82,100 km^2) and is a dimictic, oligotrophic system.[26,27] Preliminary characterization of DOM within western Lake Superior and its tributaries by UV–visible spectroscopy, FTIR, direct temperature resolved mass spectrometry and FT-ICR-MS revealed that the western arm of Lake Superior has a very different DOM composition from that found in local tributaries, which was attributed in part to the process of photo-degradation of terrestrially-derived material.[28,29] These studies showed a shift from higher molecular weight organic matter enriched in protein and lignin at the tributary sites toward compositionally different material which had a stronger aliphatic signature in the open lake. Lake Superior DOM was also shown to have a higher proportion of compounds containing N, S, and P than DOM from the tributaries.[29] Stable and radiocarbon isotopic studies show that distinct processes operating in the surface (e.g., photosynthesis) and deep waters (sediment re-suspension and pore-water intrusion) in the open lake control the relative contribution of modern and old DOM in the water column.[30] Bulk DOC across the entire lake (western through eastern basins) appears to be semilabile according to its average radiocarbon composition, which indicates an average turnover time of ≤60 years.[31]

Surface and deep-water samples were taken from seven Lake Superior sites (Fig. 1) in June 2010 when the water-column was well-mixed. The sampling sites were chosen to cover nearshore and open-lake regions, western to eastern basins in order to obtain a comprehensive view of the lake water composition. Water samples were collected via Niskin bottles from a CTD rosette at each site for DOC concentration and SPE processes.

The SPE resins applied for comparison were C18 and SDB-XC (3M Empore disks). Information on resin structure can be found on the 3M website (http://solutions.3m.com/wps/portal/3M/en_US/Empore/extraction/products/disks/product-listing/). The Empore C18 disk (pore size 60 Ǻ) contains octadecyl bonded silica sorbent which acts as non-polar stationary phase. The SDB-XC disks (pore size 80Ǻ), are made of a poly (styrenedivinylbenzene) copolymer and thus retain compounds based upon both aliphatic and aromatic interactions. Manufacturer protocols were used to obtain DOM extracts (“eR”) from 3 to 5 liters of filtered (<GF/F, nominally <0.7 μm) water samples (“init”), as shown in Fig. 2.

Briefly, each filtered sample was first acidified to pH 2 using hydrochloric acid (ACS Reagent grade). Aliquots of this acidified sample were added to the prepared SDB-XC disk (rinsed with HPLC grade acetone, pesticide grade isopropanol, Fisher Optima grade Methanol and MilliQ water) or the C18 disk (rinsed with Fisher Optima grade Methanol/MilliQ water (v/v, 9:1) one time, Fisher Optima grade Methanol two times and then MilliQ water), and then extracted under reduced pressure (around 300 mmHg). The sample on the extraction disk was then rinsed with 10ml of MilliQ water to remove salts. A 90:10(v/v) methanol:water solution was then used to recover the DOM from each sample-containing disk. To monitor SPE performances, aliquots of ‘init’ sample were analyzed by UV-Visible spectrophotometer and total organic carbon (TOC) analyzer and aliquots of extraction filtrate (‘eF’, the portion of DOM that is not retained on the resin) were also analyzed by UV-Visible spectrophotometer. The ‘eR’ samples were stored in a freezer (-20ºC), transported to lab, dried in a vacuum oven at approximately 40°C and then stored at room temperature in the dark to await analysis via UV-Visible spectrophotometer and TOC analyzer (after dissolution in known aliquots of MilliQ water), and ESI FT-ICR-MS (after addition of 70:30(v/v) methanol: water). Method blanks were also prepared by extracting 5-L of Milli-Q water in the same manner as lake water samples. All glassware used was acid-washed and baked in a muffle furnace at 450^{o}C for four hours prior to use and plastics were acid-washed in 10% HCl and rinsed with copious amounts of deionized water prior to use.

2.2. DOC analysis

DOC concentrations of whole water, ‘init’ and ‘eR’ (aliquots dried and then reconstituted in MilliQ water) at each site, both surface and deep, were determined with a Shimadzu TOC$_{VCSH}$ analyzer (Shimadzu Scientific Instruments, Inc., Columbia, MD, USA). Non-purgeable organic carbon in each sample (previously acidified to pH 2 with 6N ACS Reagent grade HCl) was analyzed after high temperature catalytic oxidation (HTCO). Potassium hydrogen phthalate (KHP) solutions were used for calibration. To assess the instrumental performance, deep sea standards (from the DOC Consensus Reference Materials Program http://yyy.rsmas.miami.edu/groups/biogeochem/CRM.html), additional KHP standards and MilliQ blanks were analyzed as samples. Three injections were performed for each sample and two more injections were performed when the standard deviation of the first three injections was greater than 2.5%; in the latter case, the closest three of the five injections were averaged to yield sample concentration.

2.3. UV/Visible spectroscopic analysis

Scanning UV-Visible spectrophotometry was performed on whole water, 'init', 'eF', and 'eR' samples with a Genesys6 scanning spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). The samples were scanned from 800 to 200 nm using a 1 cm or 5 cm quartz cuvette (depending on estimated sample concentrations). The Naperian absorption coefficient, α, was obtained using the following equation after blank correction, backscatter correction and dilution correction of absorbance:

$$\alpha_\lambda = 2.303\text{A}_\lambda * l^{-1} \qquad \text{eq. (1)}$$

where A is absorbance; λ is wavelength; and *l* is the path length (cuvette size) in meters. UV-Visible spectra of milli-Q water were used for blank correction. The average absorbance over 700-800 nm of each sample was used to correct backscatter. The sum of the volume-normalized absorption coefficients in the range of 250 nm to 400 nm was used to calculate the chromophoric dissolved organic matter (CDOM) abundance. These CDOM abundance values were then used to calculate percent recoveries of each extraction including:

$$\text{\% recovery after initial filtration} = \text{CDOM}_{\text{init}}/\text{CDOM}_{\text{whole}} \times 100\% \qquad \text{eq. (2)}$$

$$\text{\% eF recovery} = \text{CDOM}_{\text{eF}}/\text{CDOM}_{\text{init}} \times 100\% \qquad \text{eq. (3)}$$

$$\text{\% eR recovery} = \text{CDOM}_{\text{eR}}/\text{CDOM}_{\text{init}} \times 100\% \qquad \text{eq. (4)}$$

% mass balance = % eF recovery + % eR recovery eq. (5)

The following indexes were used to study the chemical properties of each sample:

E2/E3 ratio, which is the ratio of the volume normalized absorption coefficients at 250 nm (E2) and 365 nm (E3). This ratio is inversely related to molecular weight and inversely proportional to the amount of aromatic material in the sample.[32] $SUVA_{254}$, which is the ratio of the ultraviolet light absorption coefficient at wavelength 254 nm (α_{254}) to the sample DOM concentration (in mg carbon per L), giving a quantitative measure of the aromatic content per unit of organic carbon concentration.[33]

Spectral slope, S, which is determined by curve-fitting the absorption coefficient from 250 to 400 nm to a single exponential decay function using a reference wavelength of 250 nm.[34] Calculations of S were performed in Microsoft Excel, where iteration was used to maximize R^2. This index is used to give an indication of terrestrial influence and water color.[35]

As Lake Superior's water column is low in dissolved and particulate iron concentrations (0.6 to 62 nmol/L, and 22-249 nmol/L, respectively, ref. 36), no correction for iron effects upon UV-Visible absorption parameters (as described in ref. 37) was deemed necessary.

2.4. Isotopic Analyses on Bulk DOM

Radiocarbon and stable carbon isotope measurements of 'init' DOM samples were performed at the National Ocean Sciences Accelerator Mass Spectrometry Facility (NOSAMS) at the Woods Hole Oceanographic Institution as described in Zigah et al.[30] Briefly, the organic matter in <GF/F filtered water samples (acidified to pH 2)

were first oxidized to CO_2 by ultraviolet (UV) oxidation. After removal of a subsample of the purified CO_2 for δ^{13}C-DOC measurement, the remainder of the purified CO_2 was reduced to graphite and then analyzed by accelerator mass spectrometry (AMS) to determine radiocarbon content. Radiocarbon values are reported as Δ^{14}C following Stuiver and Polach. [38]

2.5. FT-ICR-MS analysis

Surface DOM samples from sites BR and WM were extracted with both C18 and SDB-XC disks and analyzed by FT-ICR-MS. Mass spectrometry was performed on a Thermo Scientific 7 Tesla Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (LTQ FT Ultra hybrid mass spectrometer) at the Woods Hole Oceanographic Institution. A linear ion trap mass spectrometer (LTQ MS) served as the ion accumulation site for high resolution analysis in the ICR cell. A mass accuracy of <1 ppm or + 0.0001 Da at 400 Da can be achieved. A scan range of 200-1100 m/z and 200 scans per analysis were used for this study. Since negative ionization results in a larger range of chemical formulae for DOM samples than positive ionization and because problematic sodiated complexes may occur in positive mode[24,39], all dried samples for this study were redissolved with 70:30(v/v) methanol:water solution and then directly injected in the ESI FT-ICR-MS system for scanning in negative mode. All sample concentration range 0.6-1.0mgC/ml before injection. Peak detection was based on a signal-to-noise ratio (S/N) greater than 5, and the peak lists were internally calibrated with a list of calibrants (Appendix Table 1). Elemental formulae were assigned to the aligned peaks (m/z values) using the

Compound Identification Algorithm (CIA) originally described by Kujawinski and Behn [24] modified in Kujawinski et al.[23] Generally C, H, O, N, P and S are considered for formulae calculation. An assignment rate of 97.9 %±1.1% was achieved for all samples. Any statistical analyses applied in this study were based on all peaks with S/N >5, rather than just the peaks with assigned elemental formulas,

2.6. Cluster Analysis and Principle Component Analysis (PCA)

The Bray-Curtis dissimilarity measurement was used to calculate the distance matrix of CA (code from Fathom toolbox, David Jones, University of Miami, http://www.rsmas.miami.edu/personal/djones/matlab/matlab.html) based on FT-ICR-MS data. This dissimilarity measurement has been found to be robust and reliable for expressing sample relationships in environmental sciences and has been widely applied to FT-ICR-MS data.[23, 29] The Bray-Curtis dissimilarity calculation only considers zero/nonzero abundance of species within the sample, thus is suitable for eliminating the influence of selective ionization suppression at different m/z caused by matrix-effects. Similarly, in applying PCA, the FT-ICR-MS data matrix consisting of only peak presence/absence information of samples from all sites, depths and season was used to minimize the influence of varying response factors. Prior to PCA, z-scoring and variance scaling was performed. PCA was also applied on UV-visible spectrometry data for all samples based on both normalized absorption coefficients from 200 to 800 cm^{-1} and the calculated spectrophotometry indices including E2/E3, SUVA254 and spectral slope. Matlab 7.0 and in-house m-files were used for data processing and multivariate analyses.

3. Results and discussion

3.1. TOC analysis, stable ($\delta^{13}C$) and radiocarbon ($\Delta^{14}C$) signatures of 'init' DOM

DOC analysis (Table 1) shows that the offshore sites and the nearshore BR site have similar DOC concentrations (range, 86.0-92.8 μM), while the nearshore ONT site exhibits higher DOC concentrations (108.6 μM), possibly due to the excess DOC input by spring snowmelt at ONT. The surface and deep samples show no significant difference in DOC concentration based on a pairwise t-test ($p = 0.6707$). Our average offshore DOC concentration (88.2 ±1.50μM) is comparable to open-lake DOM values reported by Zigah et al.[31] (88.3-94.2 μM for samples from spring and summer 2009) and lower than nearer-shore values reported by Minor and Stephens (210 ±19 μM).[40] Our average offshore DOC value is smaller than reported tributary DOC by a factor of 3 to 12 and is somewhat higher than reported surface open ocean DOC values (65-75 μM).[27,40, 41] The $\delta^{13}C$ values (-26.5±0.1‰) of all samples do not exhibit considerable variation and are close to the previously reported Lake Superior DOM $\delta^{13}C$ values for 2009 (-26~-28‰).[31] The lake's "init" DOM is primarily modern (post-1950s) at all sites according to its positive $\Delta^{14}C$ values (except ONT). An overlap of pre-aged and modern DOM from different sources may contribute to the variations of $\Delta^{14}C$ values among samples from different locations and depth.

3.2. Disk Performance Comparison Based on TOC analysis and UV/Visible Spectrometry

TOC analysis and UV-visible spectrometry were used to examine the performance of our SPE approach and to compare characteristics of ‘init’ DOM vs ‘eR’ from the SDB-XC vs C18 disks applied in this study. SPE method banks had low but measurable carbon levels. The ‘eR’ fractions of the blanks contained 2.86 μmol C and 1.67 μmol C per liter of MilliQ water that was processed through SDB-XC disk and C18 disk, respectively. These values correspond, respectively, to 13.1% and 11.0% of the average organic carbon recovered per liter of lake water. This indicates that either extraction disks or Milli-Q water introduce some contamination and that the processed blank has to be analyzed together with the samples via the same structural and compositional characterization technique (such as FT-ICR-MS in this study) to exclude the possibility of misinterpretations of sample content.

The percentage of in-situ CDOM recovered after initial filtration (<GF/F) was 94.4± 5.2%, indicating no significant contamination introduced by the initial filtration step. CDOM mass balance (eR% + eF%) averaged 94.2 ± 8.6% for SDB-XC disk SPE extraction and 85.1± 6.0% for C18 disk SPE extraction, indicating no significant CDOM contamination introduced by either resin and minimal losses through irreversible sorption. Comparison of recoveries from extractions of Lake Superior water (Table 1) showed that SDB-XC disks performed better than C18 disk in terms of both CDOM recoveries (31.0±3.80% relative to 24.4 ±3.9%) and DOC recoveries (25.8. ±9.20% relative to 18.4±6.4%). Both SDB-XC and C18 resin disks produced

higher CDOM recoveries relative to DOC recoveries at most sites (with the exception of the WM deep sample), suggesting preferential retention of CDOM relative to DOC. Our SPE DOC recoveries of Lake Superior water were generally lower than lake/river/estuary studies, [8,9,10] similar to previous Lake Superior measurements, [40,42] and comparable to or lower than marine studies.[7,9,14] This is possibly due to a combination of selective retention by solid phase resins and the organic matter composition in oligotrophic, clear-water Lake Superior.

The indices obtained via UV-visible spectrophotometry (E2/E3 ratios, $SUVA_{254}$ and spectral slope) were used to characterize the 'init' CDOM at each site, to compare XDB-XC vs C18 resin disks and to examine if the SPE substantially isolated certain portions of the DOM pool. As shown in Table 1, the E2/E3 values of 'init' water range between 8.80 to 17.5 with the general trend being higher values in offshore lake samples relative to nearshore samples (except for SM_deep). These values are somewhat higher than most of those reported for mountain lakes in Italy (4.64-11.3) with only the rockiest, above-treeline lake exhibiting an E2/E3 value in the Lake Superior range. [43] Our E2/E3 values are consistent with previous nearshore to offshore studies in Lake Superior and in a Chesapeake Bay sub-estuary, which found that the E2:E3 ratio ranged from 5.40 to 15.29 for the Superior transect and from 4.42 to 10.28 for the Chesapeake transect; in both transects the ratios increased from marsh to riverine, estuarine, coastal and open lake/ocean sites.[16] DOM from Atlantic Ocean water was found to have E2/E3 ratios of 13.5±1.6.[35] The values indicate that Lake Superior water is generally light in color and low in DOM molecular weight, with open-lake values comparable to or even higher than values from mountain lakes above the treeline and with ocean values. The E2/E3 values of recovered 'eR' fractions from both SDB-XC and C18 resin disks are lower and show less variation than the 'init' samples, indicating both extractions preferentially isolated higher molecular weight and/or more aromatic DOM. However, the SDB-XC extraction produces 'eR' E2/E3 values (8.7±0.7) closer to 'init' E2/E3 values (12.6±2.8) than does C18 extraction (7.0±0.5). This suggests that SPE with SDB-XC disk alters 'eR' molecular weight and/or aromatic DOM contents to a lesser degree than C18 extraction disk.

SUVA$_{254}$ of Lake Superior 'init' water ranges between 2.81 to 3.83 with most open-lake sample values around 3 and BR and ONT mixed season sample values at the high end of the range. Based on previous studies [33, 44] of DOM from a variety of environments (marine to dark-water rivers), Lake Superior offshore waters (SUVA$_{254}$ values ≤3 L mg^{-1} m^{-1}) contain mainly non-humic, hydrophilic and low molecular weight materials. Our SUVA$_{254}$ values are consistent with previous work in Lake Superior, where SUVA$_{254}$ values ranged from 9.0 in the St. Louis River to 2.0 in the open lake.[16] Our SUVA$_{254}$ values also show that SPE extraction with both SDB-XC and C18 resins generally leads to preferential retention of more aromatic DOM components, as suggested by the higher SUVA$_{254}$ values of 'eR' than 'init' (except for WM_deep). The SDB-XC extraction produces 'eR' SUVA $_{254}$ values (3.53±0.75) somewhat closer to 'init' SUVA $_{254}$ values (3.11±0.29) as compared to those of C18 extractions(3.60±0.67), with the exception of samples from NM and the two nearshore sites. This indicates that SPE with SDB-XC alters 'eR' aromatic DOM contents to a lesser degree than C18 extraction, however, its performance is affected by 'init' DOM composition.

For samples from all sites, and extractions with both resin types, the 'eR' fractions have lower spectral slope values than 'init', with 'eR'-C18 showing a greater deviation from 'init' water. These indicate a shift toward greater molecular size and/or aromaticity within the 'eR' fractions and are therefore consistent with the other UV-visible proxies. Thus, the combination of the UV-visible spectrometry proxy information, with $SUVA_{254}$ and E2/E3 providing insight into aromaticity and the S providing a combined view of aromaticity and molecular weight, suggests that material with higher average molecular weight and aromaticity is preferentially concentrated within the 'eR' samples by both SPE resin types, with SDB-XC SPE providing less compositional alteration and deviation from 'init' water samples.

To further investigate the compositional variation of 'eR' samples with different SPE resin types, Principle Component Analysis (PCA, ref. 45, 46) was performed on the UV-visible spectrometry data of both 'init' and 'eR' samples, using the following two data sets: A) the normalized absorption coefficients for each integer wavelength within the complete UV-visible spectrum from 200 to 800 cm^{-1} and B) calculated spectrophotometry indices, as seen in Fig. 3. Despite the possible compositional differences of 'init' DOM samples from different sites and depths, the SPE 'eR' concentrates were separated from 'init' samples, with sample preparation as the main factor driving the PCA1 variation. Fig. 3b shows that 'eR' of SDB-XC extraction may have retained more similar compositional information relative to its ‘init’ as compared to C18 extraction; this is shown by the closer spacing of SDB-XC ‘eR’ samples to the <GFF ‘init’ samples.

In summary, the combination of DOC analyses and UV/visible spectrometry parameters on bulk organic matter indicates that Lake Superior offshore water is oligotrophic, low in dissolved organic carbon content, and very clear. The DOM is low in aromatic components, with primarily non-humic, hydrophilic and low molecular weight material. In comparison of the two types of SPE disks on Lake Superior water, we found SDB-XC disks to outperform C18 disks in the isolation of DOC in terms of both higher recovery and less degree of fractionation relative to the initial DOM composition. Besides the fractionation, our blank and mass balance studies of SPE with both extraction disks show that they both introduce non-negligible contamination to ‘eR’ fractions, thus the interpretation of DOM structures based on the retentates has to be performed with care. Method blank samples have to be analyzed by whatever techniques are applied to other samples to identify blank-associated components.

3.3. FT-ICR-MS Analysis

Using extraction to isolate DOM may provide an imperfect view of the native DOM pool due to preferential recovery of certain compound types as suggested by our UV-vis spectrometry data and as reported by others. [9,13,42] Therefore, ESI FT-ICR-MS has been applied to determine molecular differences between the samples from different sites (WM vs BR) and obtained with different extraction resins (SDB-XC vs C18). All samples were analyzed under negative ion mode, which has been shown to provide greater DOM response and suppressed blank contributions relative to positive ion mode. [29,39] High abundances of DOM peaks with high intensity ranging from m/z 250 to 600 (Fig. 4) were detected, which is consistent with previous spectra from Lake Superior and its tributaries. [29] DOM peaks in samples were singly charged as determined using the isotopic distribution of carbon, [47] which is again consistent with previous DOM studies, including that of Lake Superior and selected tributaries. [29] Peaks identified in the method blank contributed 1147 to 1685 peaks to each sample, accounting for 11.8% to 15.9% of the total number of peaks in each sample, thus blank correction of the samples was performed. After blank correction, the number of peaks in samples (S/N >5) are 8691 ('eR' of SDB-XC) and 8908 ('eR' of C18) for BR surface samples and 8472 ('eR' of SDB-XC) and 8551 ('eR' of C18) for WM surface samples.

We also observed clear differences in the molecular characteristics between 'eR' of SDB-XC vs C18 extraction. The peak magnitudes of the SDB-XC samples are generally higher than those of the C18 extracted samples and different numbers of peaks can be seen within the same m/z ranges. As an example, the enlargement area of m/z 310 to 313 is shown in Fig. 4; 15 peaks exist in 'eR' of SDB-XC while only 10 peaks exist in 'eR' of C18 extracted WM surface samples.

For the same water sample (from BR surface or WM surface), the common formulae (m/z) and the unique formulae found in 'eR' obtained with different resin disks (SDB-XC vs C18) are displayed using Van Krevelen diagrams [48,49] (Fig. 5). The formulae (m/z) common to 'eR' from both extraction disks account for a very large percentage of each sample's total DOM formulae in both the offshore site (WM, 70.8% in SDB-XC 'eR' and 69.0% in C18 'eR') and nearshore site (BR, 70.3% in SDB-XC 'eR' and 69.7% in C18 'eR'), as seen in Fig. 5a and b. The prevalence of these common formulae makes sense considering that the same 'init' water sample was used for the extractions. Natural biomolecules have relatively uniform and discrete elemental compositions such that different types of molecules can be tentatively distinguished by the combination of their H/C, O/C, and N/C ratios. Seven discrete categories (indicated with Roman numbers in Fig. 5) with elemental compositions overlapping biochemical structures commonly found in soils and natural waters were defined according to H/C and O/C ratios as proposed by Hockaday et al. [48] and include: I, lipid-like compounds (O/C~0-0.2; H/C~1.5-2.3); II, peptide-like compounds (O/C~0.2-0.52; H/C~1.5-2.2); III, aminosugar-like compounds (O/C~0.52-0.7; H/C~1.5-2.2); IV, carbonhydrate-like compounds (O/C~0.7-1.2; H/C~1.5-2.4); V, condensed hydrocarbon-like compounds (O/C~0-0.25; H/C~0.5-1.2); VI, lignin-like compounds (O/C~0.25-0.67; H/C~0.7-1.5) and VII, tannin-like compounds (O/C~0.67-0.95; H/C~0.6-1.25). Although classifying species identified by FT-ICR MS into these seven categories is an obvious oversimplification, we do so here for the sake of comparing chemical variation of SPE DOM obtained at different sites and with different resins. The lignin-like fraction (within region VI, a small fraction of which contains N and P) was found to be the major component at both sites and likely derives from terrestrial inputs. The Van Krevelen space of this fraction also overlaps with the section consistent with CRAM, the most abundant identified component of DOM in the deep ocean.[50] The observation that CRAM is a major structural component (17–25%) of isolated HMW DOC in Lake Superior was reported by Zigah et al.[51] Their estimated old age (2043 BP years) for CRAM in Lake Superior HMW DOM is consistent with its refractory nature as reported for ocean systems.[50] CRAM, identified by a combination of NMR and FT-ICR-MS techniques, was also found to be a large contributor to DOM in freshwater Lake Ontario.[52] The origin of CRAM has been widely discussed; CRAM could result from degradation of cell wall and membrane components in aquatic organisms or transformation of lignin-like material originally from terrestrial sources.[50] Most of the CRAM in HMW DOC from Lake Superior is believed to be terrigenous and/or from solubilized sedimentary OC released via pore water diffusion and sediment resuspension events.[51]

The percentage of peaks unique to the SDB-XC extraction ‘eR’ and the C18 extraction ‘eR’ at both sites are similar, and range from 29.1% to 30.9% of the total number of peaks in the original sample. The compounds unique to the SDB-XC extraction ‘eR’ and the C18 extraction ‘eR’ were both found over the entire van Krevelen spaces. However, the SDB-XC extraction ‘eR’ has unique compounds that somehow follow the distribution of its common peaks, with a higher percentage of peaks in categories IV, V, VI, and VII. In contrast, the unique C18 extraction ‘eR’ compounds are more densely located within a particular region, with an H/C ratio from 1 to 2 and an O/C ratio from 0.25 to 0.5. Such dense clustering of the

C18-specific formulae in the same region of van Krevelen space has been seen before, e.g., in Perminova et al.,[12] where these authors attribute the C18 specific formulae primarily to lipids and terpenoids. Such molecular-level distribution differences can explain the UV-visible data above, where SDB-XC 'eR' samples were more similar to the 'init' water sample as compared to C18 'eR'.

3.4. Comparison of the FT-ICR-MS spectral features (elemental ratios, DBE, etc)

Elemental ratios, double bond equivalences (DBE), elemental formula percentages, and elemental percentages (Table 2) were calculated as described in Minor et al.[29] to provide an additional view of the chemical variations in samples from different sites, and extracted with different resins. The DBE (13.10-13.64) and DBE/C (0.40-0.46) values of 'eR' samples obtained with the different resins are similar. The exclusive compounds in SDB-XC 'eR' and the C18 'eR' both have higher number-averaged DBE (17.63-19.08) as compared to the total formulae in each sample; however, when DBE is normalized to the total number of C atoms in a given molecule (DBE/C), exclusive compounds in C18 'eR' have lower values (DBE/C) than the total formulae and the exclusive compounds in XC 'eR'. This is probably a function of the larger average molecular size of the C18 'eR' compounds (as shown in the UV Visible proxy data); the C18 'eR' exclusive compounds contain a higher number of rings and/or double bonds, but are, actually, a bit less unsaturated on a per C basis.

Small variations in terms of 'eR' elemental ratios are also found. The SDB-XC 'eR' compounds were found to have larger number-averaged and weight-averaged

O/C and P/C ratios, smaller H/C ratios, variable N/C and the same (very small) S/C ratios relative to the C18 'eR' compounds. For both SDB-XC and C18, the exclusive compounds have lower O/C ratios, and higher N/C ratios and somewhat higher P/C and S/C ratios. The SDB-XC exclusive compounds as compared to the C18 exclusive compounds have lower number-averaged and weight averaged H/C ratios and higher P/C ratios. That C18 retains more H-saturated components within DOM than styrene-divinylbenzene based resin has been seen before.[9]

The percentages of formulae containing S (F_S) and P (F_P) are also higher in SDB-XC 'eR' relative to C18 'eR', while the percentages of formulae containing N (F_N) are comparable for both resins. The exclusive compounds in both resin-type 'eRs' have considerably higher F_N, F_P and F_S values than the total formulae.

Combining the elemental data with the DBE numbers, the 'eR' from both SDB-XC and C18 extraction contains a large proportion of unsaturated compounds containing N, P and S, such as aromatic compounds and heteroatom-containing compounds. However, the SDB-XC resin appears to isolate DOM with slightly lower H/C ratios and slightly more double bonding; it also isolates a slightly higher percentage of formulae containing P than does the C18 resin.

Our O/C_w ratios for total formulae in the extracts (0.40~0.43) are slightly higher than previously reported values for Lake Superior C18-extracted DOM (0.38, ref 29), Lake Baikal C18-extracted DOM (0.331 to 0.348, ref 53) and C18- extracted DOM from Atlantic Ocean surface water (0.34–0.36, ref 23). Our O/C_w ratios are slightly lower than reported for PPL-extracted arctic samples from creek, river, and bay sites in the watershed of the Lena River (0.427, 0.425, and 0.445, respectively, ref 54).

H/C_w values for our samples (1.27 to 1.34) are slightly higher than previously reported for C18-extracted DOM from Lake Superior (1.22) and selected tributaries (1.08~1.28, ref 29) and similar to values for C18-extracted Lake Baikal samples (1.321-1.373, ref 53) and to values reported for creek, river, and bay PPL-extracted samples in the Lena River watershed (1.31, 1.28, and 1.23, respectively, ref 54). They are comparable to reported Dismal Swamp whole water by Sleighter and Hatcher[14] and Chesapeake Bay C18 SPE-extracted DOM (1.25 and 1.29) H/C_w values, and values reported by Kujawinski et al. for C_{18}-extracted DOM (1.30-1.32) [23] and Koch et al.[55] for PPL-extracted DOM (1.3–1.39) from ocean surface water.

Compounds containing only carbon, hydrogen and oxygen (CHO) account for 45.6% to 46.7% of all DOM formulae in each extracted sample (Table 2). This percentage is significantly lower than reported for surface ocean, deep ocean and river C18-extracted DOM samples (72.2–93.7%; ref 23), and Lake Superior tributary C18-extracted DOM samples.[29] DOM compounds with formulae of CHON account for 29.1%~32.7% (depending on sample) of all DOM formulae in each sample. These compounds are possibly in the amino-acid/protein family. The number of CHONP formulae ranges from 11.3% to 14.9% and these are likely to be nucleotides and DNA/RNA derivatives. With the above three types of compounds comprising over 95% of DOM formulae in each sample, compounds with formulae of CHOS, CHONS, CHOP and CHONSP are only minor species. Comparing 'eR' obtained with the 2 different extraction resins, SDB-XC 'eR' has a similar percentage of CHO compounds, a higher proportion of CHOS, CHONP, CHONSP compounds, a lower proportion of CHON and CHONS compounds, and similar or lower proportions of CHOP compounds. The exclusive compounds exaggerate these trends with a much lower percentage of CHO compounds and much higher percentage of compounds containing N, P and S as seen in Table 2. The majority of the exclusive formulae for both resins are CHON compounds.

Our UV-Vis data and these FT-ICR-MS data show that the use of styrene divinylbenzene-based resins vs C18 resins leads to the isolation of somewhat different portions of the DOM pool. Thus comparisons of literature values obtained with different SPE protocols should be viewed with some caution.

3.5. Comparison of the FT-ICR-MS with Multivariate analysis

Cluster Analysis (CA) and PCA were performed on the FT-ICR-MS data to further investigate the variation among samples extracted at different sites (WM vs BR) and with different resin types (SDB-XC vs C18). Cluster analysis is a convenient tool to compare samples and to visualize their differences on the basis of multidimensional data such as mass spectra. The analysis is based on the peak presence/absence of all identified molecular formulas after blank correction in the different 'eR' samples. The x-axis represents the degree of similarity with Bray-Curtis distance. As seen in Fig. 6a, the variance between samples from different sites (WM vs BR) rather than different extraction techniques drives the grouping, which is a desired result considering that we want an effective and representative isolation of DOC that represents the 'init' sample nature, which is usually affected by sampling location, season and depth. PCA was also used to explore the compositional differences and similarities among samples. As seen in the PC1 vs PC2 plot (Fig. 6b), the grouping by PCA is similar to the results from cluster analysis. PC1, which accounts for 40.70% of the total variance, separated the BR samples (on the negative side) with WM samples (on the positive side); however, at the same time, PC2, which accounts for 31.93% of the total variance, separated the C18 samples (on the positive side) with the SDB-XC samples (on the negative side). This finding confirms some retention of sample nature by SPE even with different extraction resins, and the occurring of fractionation to the initial DOM composition, affected by the type of resins. Understanding the fractionation on a molecular level will be critical to the comparison of different DOM studies of different natural systems.

4. Conclusions

Combining DOC analysis, UV-Visible spectrometry, stable and radiocarbon signatures and FT-ICR-MS allows us to assess the nature of Lake Superior DOM and to investigate and compare the performance of different SPE resins (SDVB vs C18) in disk format on the isolation of DOC in terms of both recovery and 'eR' molecular features.

DOC analyses and UV/visible spectrometry parameters show that Lake Superior offshore water is oligotrophic, low in DOC concentration, and very clear. The DOM is low in aromatic components, with primarily non-humic, hydrophilic, lower molecular weight material than in most studied fresh-water systems. The retained DOM from both extraction techniques shows a greater percentage of formulas including one or more of the heteroatoms N, S, and P and a lower percentage of CHO formulas than reported in other (ocean and river) studies.

In the comparisons of resin performance, we found that using polymeric SDVB (SDB-XC disk) leads to greater DOC and CDOM recoveries as well as less fractionation in composition relative to initial water samples. The UV-visible spectrometry proxies suggest that higher molecular weight and aromaticity material is preferentially concentrated within the ‘eR’ samples by both SPE resin types, with SDB-XC SPE providing less compositional deviation from ‘init’ water samples. Molecular-level differences among ‘eR’ obtained from the same sample using different extraction disk resins do exist based on FT-ICR-MS data, however, the shared compounds account for approximately 70% of the sample formulas. There was some preferential retention of unsaturated compounds enriched in P with the SDB-XC vs C18 resins. Comparing the two resin types, SDB-X retains similar percentages of CHO compounds, more CHOS, CHONP, CHONSP compounds, and fewer CHON, CHONS and CHOP compounds than C18. The CA and PCA analyses, which are commonly used statistical measurements of sample similarity, confirm that the greatest variance in the sample set is driven by sample location, rather than the difference in extraction techniques. However, significant fractionation relative to the initial DOM composition does occur, and the degree is affected by the type of resin. The implication of this study is that SPE retentates do contain enough chemical characteristics to represent the nature of samples, however, fractionation on a molecular level needs to be considered, especially when comparison is made between DOM samples isolated using different resins. Samples from different natural systems are also very likely to behave differently considering that the source, age and composition of their DOM are different. For oligotrophic, low-C water samples such as from Lake Superior and ocean waters, a method blank has to be considered and analyzed along with sample since considerate contamination (up to 10% in our samples) can be introduced by SPE resins.

Acknowledgments

The authors thank Elizabeth B. Kujawinski, Melissa C. Kido Soule and Krista Longnecker at the WHOI FT-MS facility for FT-ICR-MS analyses and the captain and crew of the R/V Blue Heron for sampling assistance. We thank Bruce Mattson for help with the SDB-XC SPE and the subsequent TOC and UV/Visible spectroscopic analysis. We also thank Prosper K. Zigah for assistance with sampling and for sharing radiocarbon data. The research was supported by NSF Grant OCE-0825600 (to E. C. M).

Figure Captions

Fig. 1 Lake Superior sampling sites. Stars indicate open-lake stations and circles indicate near-shore stations.

Fig. 2 Protocol used for sampling and extraction, including naming conventions followed in the text.

Fig. 3 Principle component analysis results based on A) normalized absorption coefficients from 200 to 800 cm^{-1}; B) spectrophotometry indices (E2/E3, SUVA254 and spectral slope) of all samples. Black filled squares represent <GFF 'init' samples; gray filled circles represent SDB-XC SPE DOM samples; triangles represent C18 SPE DOM samples.

Fig. 4 Negative ion mode FT-ICR-MS of a) WM surface SDB-XC SPE DOM with the enlargement of m/z 310 to 313 on the right, b) WM surface SDB-C18 SPE DOM with the enlargement of m/z 310 to 313 on the right.

Fig. 5 Van Krevelen diagram of a) formulae (m/z) found in both C18 and SDB-XC SPE DOM for surface sample at BR site; b) formulae (m/z) found in both C18 and SDB-XC SPE DOM for surface sample at WM site; c) unique compounds from SDB-XC SPE DOM at BR; d) unique compounds from SDB-XC SPE DOM at WM; e) unique compounds from C18 SPE DOM at BR; f) unique compounds from C18 SPE DOM at WM; CHO and CH compounds (black) are overlaid with CHON (blue) compounds on top of which any CHONP compounds (cyan) are placed. Any CHOP compounds (red) are on the top. The van Krevelen spaces are divided into seven discrete regions indicating by Roman numbers, modified from the diagrams proposed by Hockaday et al. [48], Kim et al. [49]; Sleighter and Hatcher [14]. The elemental ratios of compounds in the seven regions are: I, lipid-like; II, protein-like; III, aminosugar-like; IV, carbohydrate-like; V, condensed hydrocarbon-like; VI, lignin-like; VII, tannin-like. The percentage of peaks from each region relative to the total number of peaks in the sample is shown in brackets.

Fig. 6 Comparison of C18 SPE DOM and XC SPE DOM based on peak absence/presence of FT-ICR-MS. A) Cluster diagram of the samples based on original Bray-Curtis distance. The samples with lower information content (lower x-values) are more similar than those which occur farther apart; B) Principle component analysis results.

Notes and References

[a] Large Lakes Observatory and Water Resources Science Program, University of Minnesota, Duluth, MN 55812, USA. Current contact information: Earth and Ocean Sciences, University of South Carolina, 701 Sumter Street, EWS 617, Columbia, SC 29208. Phone: +0012182697516; Email: hli@geol.sc.edu

[b] Large Lakes Observatory and Department of Chemistry and Biochemistry, University of Minnesota Duluth, 2205 East 5th St, Duluth, MN 55812, USA. Tel: 218-726-7097; E-mail: eminor@d.umn.edu.

1 J. I. Hedges and J. M. Oades, Org *Geochem.*, 1997, **27**, 319 361.

2 R. Benner In: D. A. Hansell and C. A. Carlson (Eds), *Biogeochemistry of marine dissolved organic matter*, Academic Press, Boston, 2002, 69 70.

3 C. A. Carlson, H. W. Ducklow and A. F. Michaels, *Nature*, 1994, **371**, 405 408.

4 R. G. Wetzel, *Limnology: Lake and river ecosystems*, third edition. Academic Press, San Diego, 2001.

5 P. Raimbault, N. Garcia and F. Cerutti, *Biogeosciences*. 2008, **5**, 281 298.

6 T. Dittmar, R. J. Lara and G. Kattner, *Mar Chem*., 2001, **73**, 253 71.

7 E. R. M. Druffel, P. M. Williams, J. E. Bauer and J. R. Ertel, *J Geophys Res*. 1992, **97**, 15,639 15,659.

8 E. M. Thurman and R. L. Malcolm, *Environ Sci Technol*. 1981, **15**, 463 466.

9 T. Dittmar, B. Koch, H. Hertkorn and G. Kattner, *Limnol. Oceanogr. Methods*, 2008, **6**, 230 235.

10 S. Kim, A. J. Simpson, E. B. Kujawinski, M. A. Freitas and P. G. Hatcher, *Org Geochem*. 2003, **34** 1325 1335.

11 V. Roubeuf, S. Mounier and J. Y. Benaim. *Org Geochem*. 2000, **31** 127 131.

12 I. V. Perminova, I. V. Dubenkov, A. S. Konanykhin, A. I. Konstantinov, A. Y. Zherebker, M. A. Andzhushev, V. A. Lebedev, E. Bulygina, R. M. Holmes, Y. I. Kostyukevich, I. A. Popov, and E. N. Nikolaev. *Environ. Sci. Technol*. 2014, **48**, 7461 7468.

13 H. Waska, A. Koschinsky, M. J. Ruiz Chancho and T. Dittmar. Mar Chem. 2015, **173**, 78 92.

14 R. L. Sleighter, P. G. Hatcher, *Mar Chem*., 2008, **110**, 140 152.

15 R. Benner, B. Biddanda, B. Black and M. McCarthy, *Mar Chem*., 1997, **57**, 243 263.

16 B. M. Stephens and E. C. Minor, *Aquat Sci*., 2010, **72**, 403 417.

17 T. A. Vetter, E. M. Perdue, E. Ingall, J-F. Koprivnjak and Ph. Pfromm, *Sep Purif Technol*., 2007, **56**, 383 387.

18 J-F. Koprivnjak, Ph. Pfromm, E. Ingall, V. A. Vetter, P. Schmitt-Koplin, N. Hertkorn, M. Frommberger, H. Knicker, and E. M. Perdue, *Geochim Cosmochim Acta*., 2009, **73**, 4215 4231.

19 E. B. Kujawinski, *Environ Forensics*., 2002, **3**, 207 216.

20 R. N. Mead, K. M. Mullaugh, G. Brooks Avery, R. J. Kieber, J. D. Willey and D. C. Podgorski, *Atmos Chem Phys*., 2013, **13**, 4829 4838.

21 A. Nebbioso and A Piccolo, *Anal Bioanal Chem*., 2013, **405**, 109 214.

22 A. M. Kellerman, T. Dittmar, D. N. Kothawala and L. J. Tranvik. Nat Commun., 2014, DOI: 10.1038/ncomms4804.

23 E. B. Kujawinski, K. Longnecker, N. V. Blough, R. Del Vecchio, L. Finlay, J. B. Kitner and S. J. Giovannoni, *Geochim Cosmochim Acta* 2009, **73**, 4384 4399.

24 E. B. Kujawinski and M. D. Behn, *Anal Chem*., 2006, **78**, 4363 4373.

25 H. A. N. Abdulla, P. L. Sleighter and P. G. Hatcher, *Anal Chem*., 2013, **85**, 3895 3902.

26 C. E. Herdendorf, *J Great Lakes Res*., 1982, **3**, 379 412.

27 N. R. Urban, M. T. Auer, S. A. Green, X. Lu, D. S. Apul, K. D. Powell and L. Bub, *J Geophys Res*., 2005, **110**, C06S90, doi:10.1029/2003JC002230.

28 E. C. Minor, J. J. Boon, H. R. Harvey and A. Mannino, *Geochim Cosmochim Acta*. 2001, **65**, 2819 2834.

29 E. C. Minor, C. J. Steinbring, K. Longnecker and E. B. Kujawinski, *Org Geochem*., 2012, **43**, 1 11.

30 P. K. Zigah, E. C. Minor, J. P. Werne and S. L. McCallister, *Limnol Oceanogr*., 2011, **56**, 867 886.

31 P. K. Zigah, E. C. Minor and J. P. Werne, *Global Biogeochem Cycles*, 2012, doi:10.1029/2011GB004132.

32 J. Peuravuori and K. Pihlaja, *Anal Chim Acta*., 1997, **337**, 133 149.

33 J. L. Weishaar, G. R. Aiken, B. A. Bergamaschi, M. S. Fram, R. Fujii and K. Mopper, *Environ Sci Technol*., 2003, **37**, 4702 4708.

34 C. A. Stedmon, S. Markager and H. Kaas, *Estuar Coast Shelf Sci*., 2000, **51**, 267

278.

35 J. R. Helms, A. Stubbins, J. D. Ritchie, E. C. Minor, D. J. Kieber and K. Mopper, *Limnol Oceanogr*., 2008, **53**, 955 969.

36 C. S. Hassler, S. M. Havens, G. S. Bullerjahn, R. M. L. McKay and M. R. Twiss, *Limnol. Oceanogr*., 2009, **54**, 987 1001.

37 B. A. Poulin, J. N. Ryan and G. R. Aiken, *Environ. Sci. Technol*., 2014, **48**, 10098 10106.

38 M. Stuiver and H. A. Polach, *Radiocarbon*, 1977, **19**, 355 363.

39 N. Hertkorn, M. Frommberger, M. Witt, B. P. Koch, Ph. Schmitt-Kopplin and E. M. Perdue, *Anal Chem*., 2008, **80**, 8908 8919.

40 E. C. Minor and B. Stephens, *Org Geochem*., 2008, 3**9**, 1489 1501.

41 D. A. Hansell and C. A. Carlson, *Oceanography*, 2001, **14**, 41 49.

42 B. R. Kruger, B. J. Dalzell and E. C. Minor, Aquat Sci., 2011, 73, 405 417.

43 E. D. Laurentiis, M. Minella, V. Maurino, C. Minero, M. Brigante, G. Mailhot and D. Vione, Chemosphere, 2012, **88**, 1208 1213.

44 R. G. M. Spencer, K. D. Butler and G. R. Aiken, J. Geophys. Res., 2012, **117**, G03001.

45 O. M. Kvalheim, D. W. Aksnes, T. Brekke, M. O. Eide and E. Sletten, *Anal Chem*., 1985, **57**, 2858 2864.

46 O. D. Sanni, M. S. Wagner, D. Briggs, D. G. Castner and J. C. Vickerman, *Surf Interface Anal*., 2002, **33**, 715 728.

47 A. C. Stenson, W. M. Landing, A. G. Marshall and W. T. Cooper, *Analyt Chem.,*

2002, **74**, 4397 4409.

48 W. C. Hockaday, J. M. Purcell, A. G. Marshall, J. A. Baldock and P. G. Hatcher, Limnol. Oceanogr.: Methods, 2009, 7 81 95.

49 S. Kim, R.W. Kramer and P.G. Hatcher, *Analyt. Chem*., 2003, **75**, 5336 5344.

50 N, Hertkorn, R. Benner, M. Frommberger, Ph. Schmitt-Kopplin, M. Witt, K. Kaiser, A. Kettrup and J. I. Hedges, *Geochim Cosmochim Acta*, 2006, **70**, 2990 3010.

51 P. K. Zigah, E. C. Minor, H. A. N. Abdulla, J. P. Werne and P. G. Hatcher, *Geochim Cosmochim Acta*, 2014, **127**, 264 284.

52 B. Lam, A. Baer, M. Alaee, B. Lefebvre, A. Moser, A. Williams and A. J. Simpson, *Environ Sci Technol*., 2007, **41**, 8240 8247.

53 Y. Sugiyama, P. G. Hatcher, R. L. Sleighter, T. Suzuki, C. Wada, T. Kumagai, O. Mitamura, T. Katano, S. Nakano, Y. Tanaka, V. V. Drucker ,V. A. Fialkov and M. Sugiyama, Limnol., 2014, **151**, 127 139.

54 I. Dubinenkov, R. Flerus, P. Schmitt-Kopplin, G. Kattner and B. P. Koch, Biogeochem., 2015, **123**, 1 14.

55 B. P. Koch, K.-U. Ludwichowski, G. Kattner, T. Dittmar and M. Witt, *Mar. Chem*., 2008, **111**, 233 241.

Table 1

Sampling information, DOC concentration, stable carbon ($\delta^{13}C$) and radiocarbon signatures of ‘init’ DOM, CDOM and DOC recoveries (% CDOM and % DOC of ‘eR’ to ‘init’ samples) and UV-Visible spectrophotometry indices for ‘init’ and ‘eR’ samples of both C18 and SDB-XC extractions.

Sample		CM	CM	EM	EM	NM	NM	SM	SM	WM	WM	BR	ONT
Water depth m		258	258	248	240	213	216	398	386	171	171	19	20
Sample depth m		5	190	5	210	5	150	5	340	5	127	4	5
‘init’ DOC µM		86.5	89.5	89.8	87.8	88.2	87.8	87.8	86.0	87.8	91.0	92.8	108.6
‘init’ $\delta^{13}C$ ‰*		--	--	-26.1	-25.9	--	--	-25.9	-26.0	-26.0	-29.0	--	-28.2
‘init’ $\Delta^{14}C$ ‰*		--	--	44.6	49.5	--	--	40.0	47.5	57.4	156.9	--	-18.5
CDOM Recovery %	XC	30.9	25.5	28.0	31.6	31.4	34.0	30.8	24.0	35.6	34.2	28.8	36.4
	C18	19.0	27.0	24.2	26.2	26.6	24.2	22.5	19.9	19.1	28.5	23.7	31.8
DOC Recovery %	XC	23.1	22.9	21.6	22.5	22.2	22.3	22.2	23.1	23.2	54.8	26.0	25.7
	C18	11.8	20.1	15.8	17.5	18.2	17.7	14.7	17.1	11.0	36.3	21.3	19.6
E2/E3	‘init’	15.5	12.2	12.2	17.5	12.3	12.7	10.2	8.80	15.5	15.2	9.08	9.91
	XC ‘eR’	8.91	9.40	9.40	8.98	8.91	8.29	9.08	9.10	8.77	8.57	7.55	7.21
	C18 'eR'	7.13	7.24	7.25	7.28	7.28	7.41	7.44	7.13	7.12	6.61	6.14	5.82
SUVA254 L/(m*mg)	‘init’	3.02	2.98	2.81	2.83	3.00	3.02	2.90	3.12	3.16	3.10	3.50	3.83
	XC ‘eR’	3.55	3.01	3.38	3.47	3.81	4.07	3.73	3.26	4.02	1.66	3.63	4.83
	C18 'eR'	3.89	3.36	3.59	3.41	3.65	3.50	3.82	3.34	4.32	1.92	3.62	4.72
Spectral slope	‘init’	0.022	0.022	0.022	0.023	0.022	0.021	0.021	0.022	0.022	0.022	0.019	0.019
	XC ‘eR’	0.019	0.019	0.019	0.019	0.019	0.018	0.019	0.019	0.019	0.018	0.017	0.017
	C18 'eR'	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.016	0.014

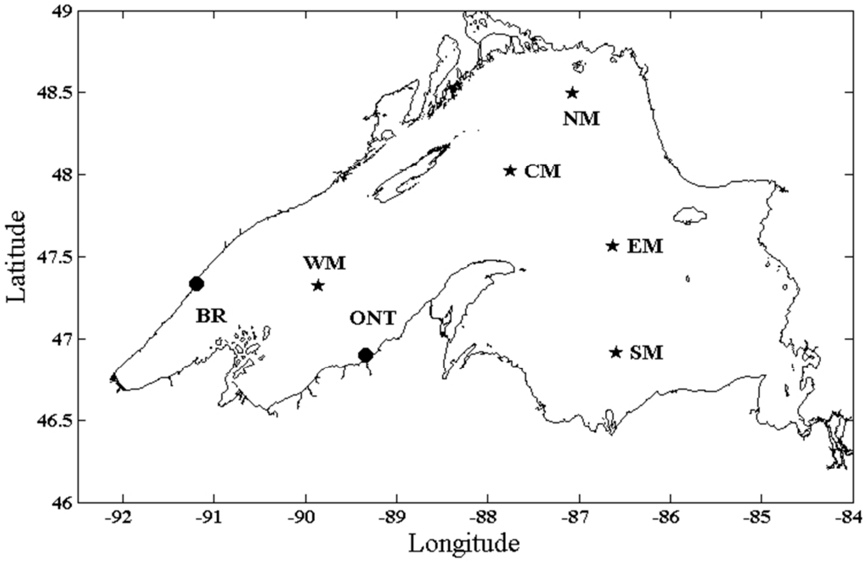
* From Zigah 2012; -- No measurements

Table 2

Elemental data from formula assignments. The subscript w indicates magnitude-averaged values; the subscript n indicates number-averaged values as described in Minor et al. 2012.

	BR surf XC	BR surf C18	BR surf XC exclusive	BR surf C18 exclusive	WM surf XC	WM surf C18	WM surf XC exclusive	WM surf C18 exclusive
DBE[*]	13.22	13.10	18.44	17.63	13.40	13.64	18.46	19.08
DBE/C	0.42	0.40	0.43	0.36	0.45	0.44	0.46	0.40
H/C_n	1.314	1.353	1.263	1.394	1.273	1.304	1.206	1.309
O/C_n	0.413	0.398	0.386	0.338	0.419	0.412	0.381	0.361
N/C_n	0.156	0.159	0.222	0.226	0.166	0.162	0.222	0.207
P/C_n	0.009	0.007	0.015	0.009	0.013	0.009	0.022	0.010
S/C_n	0.003	0.003	0.007	0.005	0.003	0.003	0.008	0.006
H/C_w	1.299	1.343	1.271	1.418	1.273	1.298	1.214	1.311
O/C_w	0.431	0.403	0.387	0.335	0.424	0.419	0.366	0.360
N/C_w	0.094	0.106	0.219	0.230	0.128	0.106	0.231	0.205
P/C_w	0.006	0.005	0.015	0.009	0.013	0.007	0.028	0.012
S/C_w	0.002	0.002	0.007	0.005	0.001	0.001	0.008	0.005
FN	47.50	48.28	67.35	68.32	48.78	48.29	66.43	64.25
FP	14.68	13.28	20.36	15.38	17.98	14.97	26.53	16.35
FS	4.96	4.58	9.93	8.31	4.65	3.70	11.18	7.83
%CHO	46.15	45.60	18.27	18.69	45.83	46.66	19.73	23.25
%CHON	30.79	32.71	43.72	48.91	29.13	31.76	36.44	44.89
%CHOS	1.55	1.22	2.80	1.63	0.89	0.60	1.87	0.89
%CHONS	2.30	2.49	4.45	4.90	2.12	2.20	4.85	5.01
%CHOP	0.62	0.63	1.46	1.42	0.45	0.73	0.95	1.85
%CHONP	12.58	11.36	15.56	11.39	14.93	12.82	19.21	12.11
%CHONSP	0.55	0.36	1.42	0.73	0.76	0.36	2.07	0.73

[*] The equation $DBE=(C+{}^{13}C)-H/2+N/2+1$ was used to calculate the DBE of each molecule. Then these DBEs were number-averaged, i.e., were calculated by dividing the sum of the DBEs by the total number of assigned formulae, for the value reported here.



230x139mm (96 x 96 DPI)

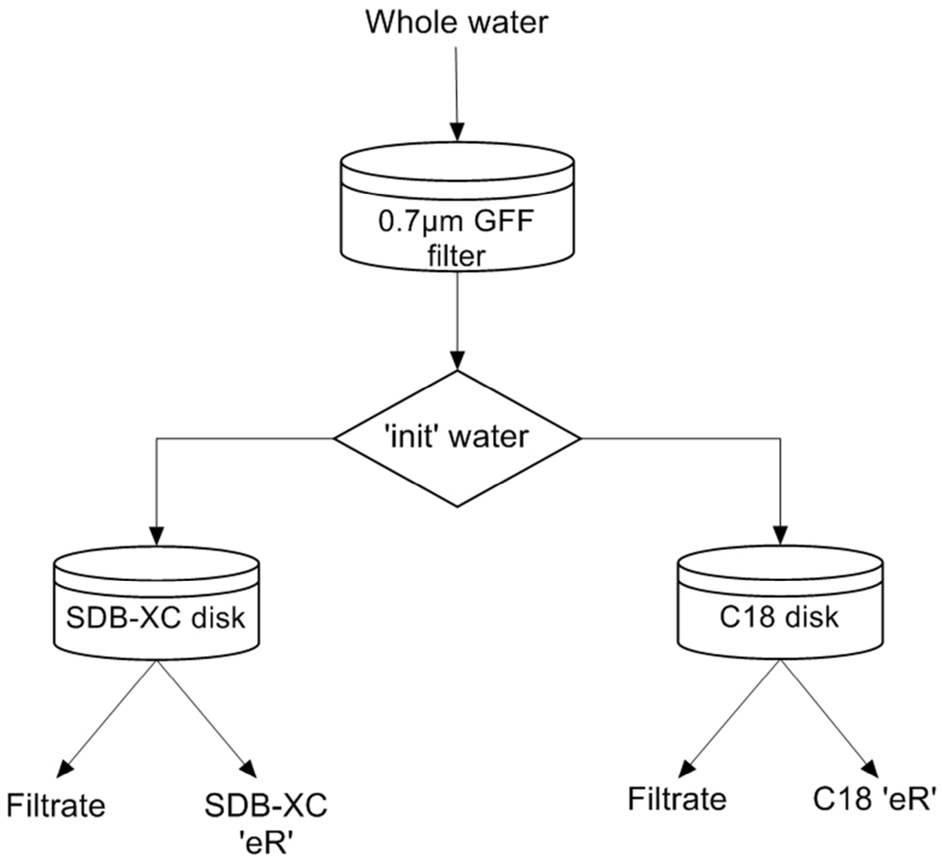


Fig. 2 Protocol used for sampling and extraction, including naming conventions followed in the text.

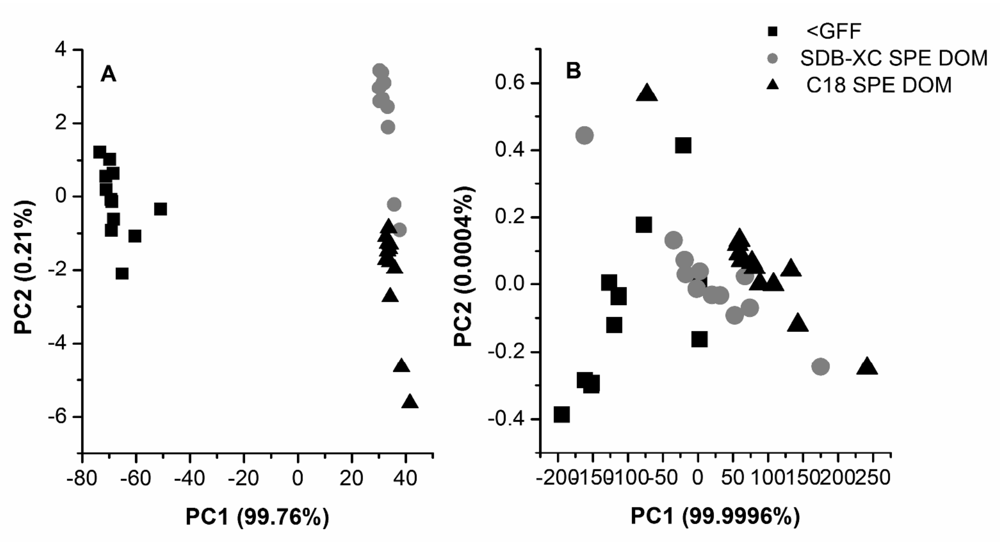
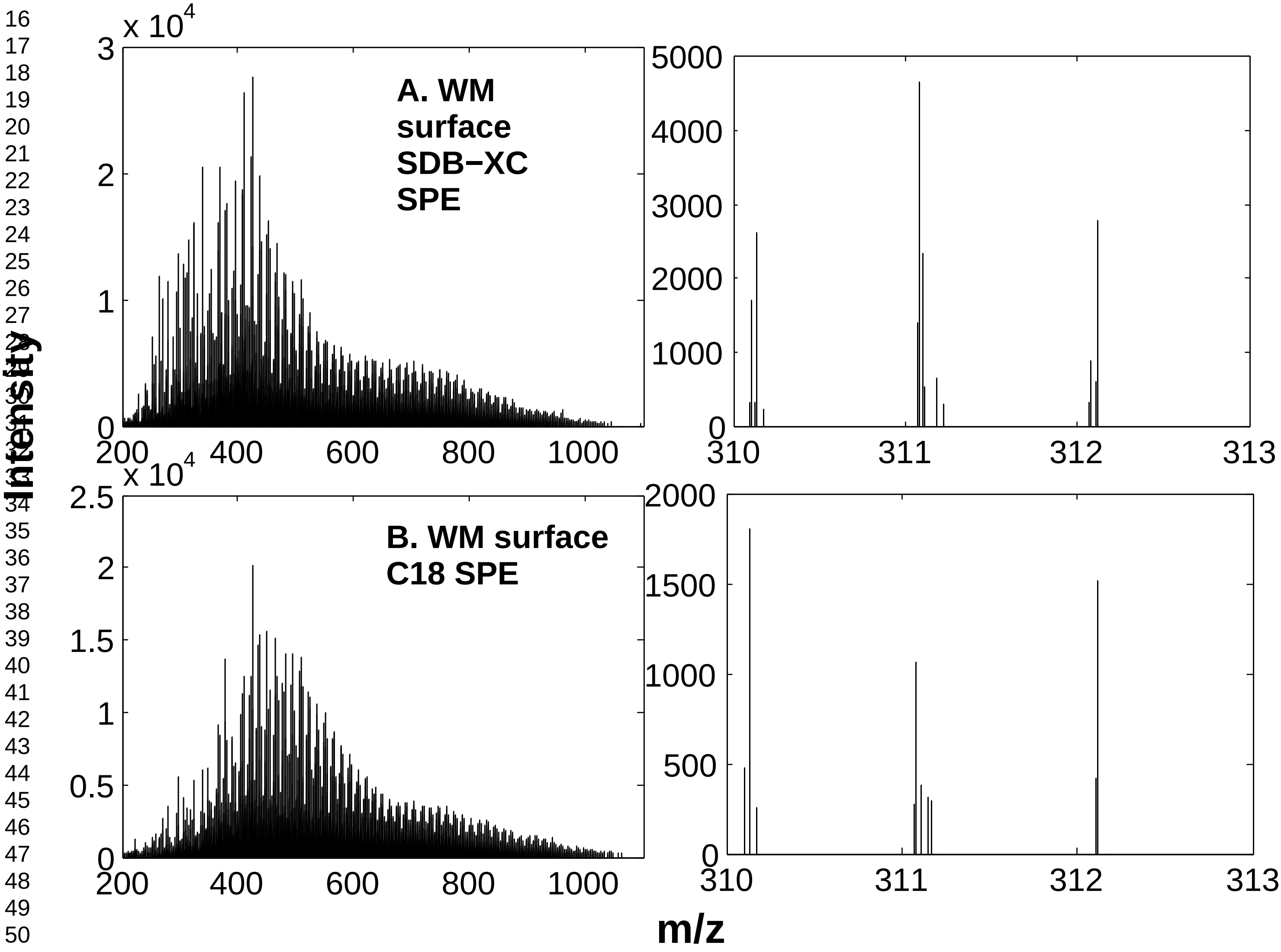
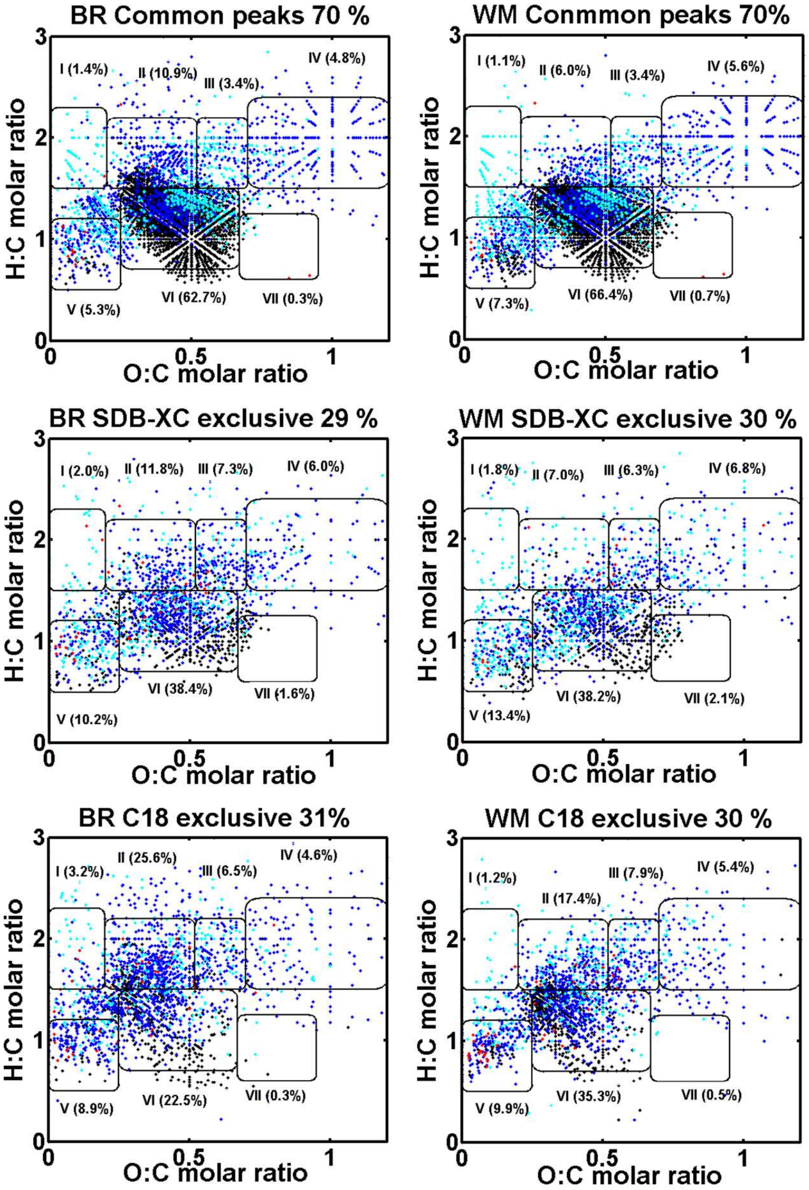


Fig. 3 Principle component analysis results based on A) normalized absorption coefficients from 200 to 800 cm-1; B) spectrophotometry indices (E2/E3, SUVA254 and spectral slope) of all samples. Black filled squares represent <GFF 'init' samples; gray filled circles represent SDB-XC SPE DOM samples; triangles represent C18 SPE DOM samples.
660x355mm (150 x 150 DPI)


x 10^4
3
2
1
0
200 400 600 800 1000
A. WM surface SDB−XC SPE
5000
4000
3000
2000
1000
0
310 311 312 313
x 10^4
2.5
2
1.5
1
0.5
0
200 400 600 800 1000
B. WM surface C18 SPE
2000
1500
1000
500
0
310 311 312 313
Intensity
m/z



211x308mm (144 x 144 DPI)

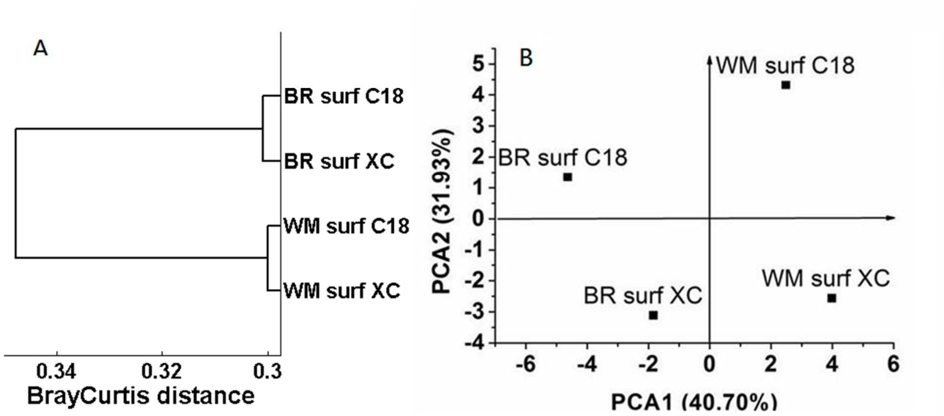


Fig. 6 Comparison of C18 SPE DOM and XC SPE DOM based on peak absence/presence of FT-ICR-MS. A) Cluster diagram of the samples based on original Bray-Curtis distance. The samples with lower information content (lower x-values) are more similar than those which occur farther apart; B) Principle component analysis results.
272x118mm (96 x 96 DPI)