

# Investigating the Ionization of Dissolved Organic Matter by Electrospray

Claudia Patriarca, Andrea Balderrama, Martina Može, Per J. R. Sjöberg, Jonas Bergquist, Lars J. Tranvik, and Jeffrey A. Hawkes\*



Cite This: *Anal. Chem.* 2020, 92, 14210–14218



Read Online

ACCESS |



Metrics & More

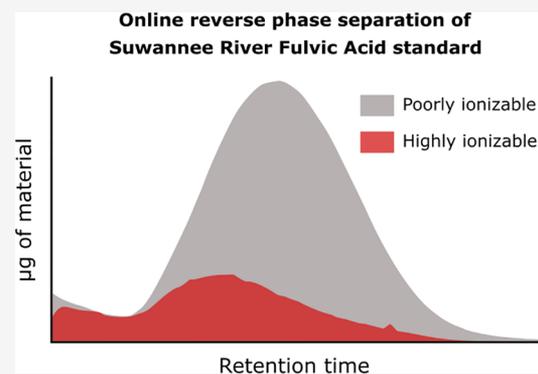


Article Recommendations



Supporting Information

**ABSTRACT:** Electrospray ionization (ESI) operating in the negative mode coupled to high-resolution mass spectrometry is the most popular technique for the characterization of dissolved organic matter (DOM). The vast molecular heterogeneity and the functional group diversity of this complex mixture prevents the efficient ionization of the organic material by a single ionization source, so the presence of uncharacterized material is unavoidable. The extent of this poorly ionizable pool of carbon is unknown, is presumably variable between samples, and can only be assessed by the combination of analysis with a uniform detection method. Charged aerosol detection (CAD), whose response is proportional to the amount of nonvolatile material and is independent from the physicochemical properties of the analytes, is a suitable candidate. In this study, a fulvic acid mixture was fractionated and analyzed by high-pressure liquid chromatography–mass spectrometry in order to investigate the polarity and size distributions of highly and poorly ionizable material in the sample. Additionally, DOM samples of terrestrial and marine origins were analyzed to evaluate the variability of these pools across the land–sea aquatic continuum. The relative response factor values indicated that highly ionizable components of aquatic DOM mixtures are more hydrophilic and have lower molecular weight than poorly ionizable components. Additionally, a discrepancy between the samples of terrestrial and marine origins was found, indicating that marine samples are better represented by ESI than terrestrial samples, which have an abundant portion of hydrophobic poorly ionizable material.



## INTRODUCTION

Dissolved organic matter (DOM) from aquatic environments is a complex and heterogeneous mixture characterized by material with a wide range of molecular weight (MW), polarity, structure, and stability. Its extreme isomeric complexity prevents the complete isolation of individual species, even using high-resolution mass spectrometry (HRMS) coupled with chromatographic separation.<sup>1–4</sup>

No matter which ionization source is selected prior to mass spectrometry analysis, ionization is not uniform.<sup>1,5,6</sup> The intrinsic complexity of natural aquatic mixtures implies that only a part of the DOM pool is in fact ionizable, and different compounds have different sensitivities to ionization. Moreover, in complex mixtures the effects of ionization suppression and enhancement become difficult to predict. Most of the signals observed in negative electrospray ionization (ESI(−)) coupled to HRMS are carboxylic acids,<sup>7</sup> probably due to their abundance and tendency to easily ionize.<sup>8,9</sup> Other compound types (lignin polymers, polysaccharides, and proteins)<sup>10,11</sup> are characterized by lower or null ionization efficiency and are therefore underrepresented in HRMS studies.

While NMR is able to measure the proton environment of all dissolved organic molecules, it requires a high concentration and

is not able to resolve different species in the complex mixture. Among the techniques sensitive enough to detect low concentrations, none are able to detect all the species constituting DOM; therefore, a combination of multiple techniques is advisable in order to gain more insights into the composition of the natural material. The combination of qualitative and quantitative techniques is promising because it allows contextualization of physicochemical properties of the material according to its abundance.<sup>5,12–14</sup> In the absence of a reference standard, as in the analysis of natural complex mixtures, quantitative information can be achieved only by employing a universal detector.<sup>14</sup> A technique able to unambiguously detect all nonvolatile material in a sample and achieve a near uniform response is charged aerosol detection (CAD). The detector response is independent from the physicochemical properties of the material and its signal is

Received: August 13, 2020

Accepted: September 17, 2020

Published: September 17, 2020

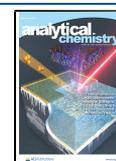


Table 1. Acids Tested on the Reverse-Phase System Coupled to ESI(-)-MS and CAD<sup>a,b</sup>

name	exact mass	formula	number of carboxylic acid groups	log <i>P</i>	relative response factor alone (TAC/ $\mu$ g) $\times 10^6$	relative response factor diluted in SRFA (TAC/ $\mu$ g) $\times 10^6$	relative enhancement in SRFA
1,4-cyclohexanedicarboxylic acid	172.074	C <sub>8</sub> H <sub>12</sub> O <sub>4</sub>	2	0.5	2.69 $\pm$ 0.16	1.12 $\pm$ 0.16	0.41
hippuric acid	179.058	C <sub>9</sub> H <sub>9</sub> NO <sub>3</sub>	1	0.3	3.11 $\pm$ 0.19	3.18 $\pm$ 0.27	1.02
aspirin	180.042	C <sub>9</sub> H <sub>8</sub> O <sub>4</sub>	1	1.2	2.09 $\pm$ 0.13	0.00 $\pm$ 0.00	0.00
cyclohexylsuccinic acid	200.105	C <sub>10</sub> H <sub>16</sub> O <sub>4</sub>	2	2	5.11 $\pm$ 0.21	16.32 $\pm$ 1.39	3.20
ibuprofen	206.131	C <sub>13</sub> H <sub>18</sub> O <sub>2</sub>	1	3.5	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	0.00
(5-formyl-2-methoxyphenoxy) acetic acid	210.053	C <sub>10</sub> H <sub>10</sub> O <sub>5</sub>	1	0.4	3.59 $\pm$ 0.22	3.33 $\pm$ 0.31	0.93
sinapic acid	224.068	C <sub>11</sub> H <sub>12</sub> O <sub>5</sub>	1	1.5	1.61 $\pm$ 0.11	0.85 $\pm$ 0.07	0.53
2,5-dihydroxy-1,4-benzenediacetic acid	226.048	C <sub>10</sub> H <sub>10</sub> O <sub>6</sub>	2	0.1	3.17 $\pm$ 0.08	1.62 $\pm$ 0.12	0.51
3-ethyl-3-phenyl glutaric acid	236.105	C <sub>13</sub> H <sub>16</sub> O <sub>4</sub>	2	2.1	4.88 $\pm$ 0.22	9.77 $\pm$ 0.77	2.00
benzene-1,3,5-triacetic acid	252.22	C <sub>12</sub> H <sub>12</sub> O <sub>6</sub>	3	0.3	2.74 $\pm$ 0.23	0.03 $\pm$ 0.03	0.01
2-(4-(2,2-dicarboxy-ethyl)-2,5-dimethoxy-benzyl-malonic) acid	370.09	C <sub>16</sub> H <sub>18</sub> O <sub>10</sub>	4	1	3.98 $\pm$ 0.39	5.24 $\pm$ 0.69	1.32
fusidic acid	516.345	C <sub>31</sub> H <sub>48</sub> O <sub>6</sub>	1	5.5	1.86 $\pm$ 0.16	3.30 $\pm$ 0.53	1.78
carboxolone	570.356	C <sub>34</sub> H <sub>50</sub> O <sub>7</sub>	2	6.4	0.99 $\pm$ 0.07	3.27 $\pm$ 0.60	3.31
glycyrrhizic acid	822.404	C <sub>42</sub> H <sub>62</sub> O <sub>16</sub>	1	3.7	1.13 $\pm$ 0.08	0.00 $\pm$ 0.00	0.00
median					2.7	2.4	
mean					2.6	3.4	
standard deviation					1.5	4.6	
standard error					0.4	1.2	

<sup>a</sup>In addition to the exact mass, chemical formula, and number of carboxylic acid groups, the log *P* calculated by XLogP3 3.0 (PubChem release 2019.06.18) is given. The result for the relative ESI(-)-MS response factor is displayed as counts  $\times$  time/ $\mu$ g injected, both in clean solution and after dilution in SRFA. The relative enhancement in relative response factor after dilution is shown as a ratio in the final column. <sup>b</sup>The error stated for the relative response factors is the uncertainty of the slope at 95% confidence level.

proportional only to the amount of material transferred to the analyzer.<sup>15–17</sup> This analyzer is compatible with simultaneous analysis by ESI-HRMS because of similar requirements for solvent volatility.

In this study, molecular information provided by HRMS analysis, and limited by the selectivity of the ESI source employed to generate gas-phase ions, was merged with the quantitative response from the CAD. Using this strategy, the abundance of the material that is highly and poorly ionizable in the set ionization conditions was investigated in complex aquatic DOM samples. The material was assessed based on its relative response factor, defined as the ratio between the mass spectrometric signal (MS) and the amount of the material ( $\mu$ g) eluting from the chromatographic system in a given time window and the data were compared with a set of purchased carboxylic acids.

## EXPERIMENTAL SECTION

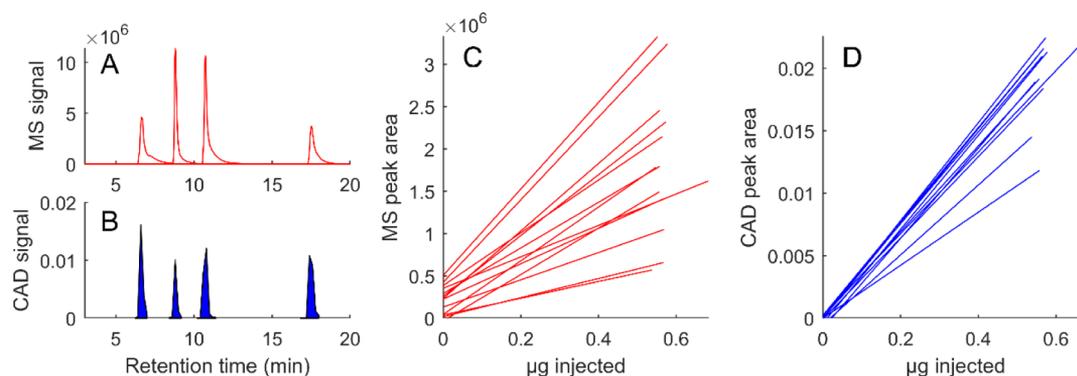
**Samples. Carboxylic Acids.** A group of 14 carboxylic acids was purchased from Sigma-Aldrich with a wide range of polarities (log *P* 0.1–6.4) and molecular masses (172–822 Da; Table 1), see Table S12 for further details. The acids were accurately weighed into three mixtures with well-resolved retention times at approximately 10 mg/L each.

**Aquatic Samples.** Four powdered reference samples were obtained from the International Humic Substances Society: Suwannee River fulvic acid (SRFA), Pony Lake fulvic acid (PLFA), Nordic NOM (NNOM), and Elliott Soil Fulvic Acid (ESFA). Two samples were collected from the river in Munkedal (river; 58.4615 N, 11.6857 E) and the fjord Gullmarn (fjord; 58.3192 N, 11.5411 E) in western Sweden. All samples were prepared by solid-phase extraction (SPE), following the method outlined in Dittmar et al., 2008,<sup>18</sup> but for the river and fjord samples with acetonitrile as the elution solvent. An additional sample of un-extracted SRFA was prepared for

fractionation with high pressure size exclusion chromatography (HPSEC).

**Size Fractionation of SRFA.** SRFA was fractionated by high-pressure size exclusion chromatography (HPSEC). The high-pressure liquid chromatographic (HPLC) system was an Agilent 1100 equipped with a Tosoh column (TSKGel GW3000SW 300  $\times$  7.5 mm, 10  $\mu$ m) operating in isocratic mode with mobile phase 20% methanol, 25 mM ammonium acetate. Four fractions along with a mobile phase blank were collected, and in order to isolate enough material for the reversed phase analysis, the sample was injected (40  $\mu$ L) and fractionated multiple times. The abundance of the isolated material was evaluated on the basis of the CAD signal in a previous run and ultimately, different volumes were isolated for each fraction (fraction 1: 4  $\times$  2 mL; fraction 2, 3 and the blank: 4  $\times$  1 mL; fraction 4: 4  $\times$  2.5 mL). The mobile phase was not optimized for the reverse-phase separation and mass spectrometric detection, so it was removed by a rotary evaporator (Rotavapor, Buchi; water bath 45  $^{\circ}$ C); the dried fractions were recovered with 1 mL of methanol, which was further evaporated under a gentle stream of nitrogen gas and finally re-dissolved in 300  $\mu$ L of 5% acetonitrile solution suitable for the reverse-phase analysis.

**Separation and Detection.** Samples and HPSEC fractions were analyzed by HPLC-CAD-ESI-MS. The HPLC system was an Agilent 1100 equipped with a Polar-C18 column (Phenomenex, Kinetex Polar-C18 100  $\times$  2.1 mm, 2.6  $\mu$ m) with mobile phases 0.1% formic acid (A) and 80% acetonitrile and 0.1% formic acid (B) pumped at 220  $\mu$ L/min. The gradient was programmed to remain isocratic for 1.5 min at 1% B, ramp to 99% B at 20 min, remain isocratic for 3 min, ramp back to 1% B at 23.2 min, and re-equilibrate at 1% B until 30 min. In gradient elution, the continuous change in the mobile phase composition affects the nebulization process (surface tension and desolvation) altering the CAD and ESI response. To overcome this effect, a compensation gradient ensuring a constant mobile



**Figure 1.** CAD and ESI-MS response of purchased carboxylic acids: (A) MS profile of four selected acids; (B) CAD signal of four selected acids; and (C,D) calibrations curves of tested acids detected by MS (red,  $n = 13$ ) and CAD (blue,  $n = 11$ ). Only the regression lines and not the calibration points are shown, all were highly linear and well fitted ( $r^2 > 0.98$ ). Calibration points and  $r^2$  values for each acid are shown in Figures S11 and S12.

phase composition during the analysis was applied.<sup>19</sup> A counter gradient was provided by a second pump via a T-piece, accounting for column volume, to make the overall solvent composition post-column isocratic 40% acetonitrile and 0.1% formic acid. The eluting material (440  $\mu\text{L}/\text{min}$ ) was split between the charged aerosol detector (Corona RS, Thermo Fisher) and electrospray ionization mass spectrometer (Orbitrap Velos, Thermo Fisher), with the majority of flow ( $\sim 80\%$ ) diverted to the charged aerosol detector with a wider bore tubing.

For the carboxylic acid mixtures (10 ppm), volumes of 0, 10, 20, 40, and 60  $\mu\text{L}$  were injected for abundances of 0, 100, 200, 400, and 600 ng. In addition to the clean calibration curves, the carboxylic acid mixtures were also diluted into SRFA (1080 ppm) in 5% acetonitrile at different concentrations—0, 50, 100, 200, and 300 ppb, in order to produce calibration curves in lower concentrations with ionization competition. The injection volume was set to 20  $\mu\text{L}$  in order to inject 0, 1, 2, 4, and 6 ng of the compound in the presence of a total SRFA amount of 20  $\mu\text{g}$ . The complex sample was prepared to 1000 ppm substance for the powders and SPE extracts (SRFA, PLFA, ESFA and NNOM, river, fjord) and injected at 20  $\mu\text{L}$  (20  $\mu\text{g}$  injected).

The CAD instrument was connected to a flow of pure nitrogen gas at 35 psi and the detector was set to range 20 pA with the output connected to the analogue 1 V input of the Orbitrap. The Orbitrap ESI was set to 3 kV in negative mode, capillary temperature 275  $^{\circ}\text{C}$ , and sheath gas 28, all other settings can be found in the Supporting Information (Table S11). The ion optic settings were tuned using SRFA in direct infusion mode by auto-optimizing the signal at  $m/z$  369 in ion trap mode.

**Data Processing. CAD.** Data from standard carboxylic acids and samples were processed with the same method in MATLAB. The signal from the CAD was zero-baseline corrected using the data between 22 and 24 min, then the data were binned by averaging the signal in 0.2 min bins and multiplying the result by the bin width (0.2), in order to obtain an approximation of the area of the signal in the bin. The binned signals were summed to calculate the total signal area. Complex samples were quantified by the calibration of the obtained bin areas with the average regression slope of standard carboxylic acids (Figure S11).

**ESI-MS.** The ions obtained after ESI of the standard carboxylic acids were inspected for each acid and ions (singly charged, doubly charged, adducts, and dimers) can be found in the Table S12. In each case, all detected ions were summed to calculate the total relative response factor of the acid. The

summed current of the ions obtained was binned to a 0.2 min window by averaging the intensity in each bin and multiplying by 0.2 to obtain the assigned intensity area for each acid (Figure S12). These were calculated at four injected amounts in triplicate to determine the relative response factor (the slope) for each acid. The evaluated intercepts were often non-zero (Figure S12), indicating some degree of curvature at low abundance. These processing steps were the same for the analysis in which the acids were diluted in SRFA, for which the response was also linear.

For complex samples, peak lists were extracted for each transient from mzXML files in MATLAB. Noise was evaluated as the median of signal intensity with mass defect 0.3–0.5 in each transient. Signals that were  $>3\times$  noise were considered for formula assignment. Formulas were assigned using a theoretical formula list with the following constraints: C 4–40, H 4–80, O 0–35, N 0–1,  $^{13}\text{C}$  0–1, S 0–1,  $m/z$  150–700, H/C 0.3–2.4, O/C 0–1, valence neutral, single negative charge,  $N + S + ^{13}\text{C} \leq 1$ , and double bond equivalence minus oxygen  $\leq 10$ , leaving a theoretical formula list with 28,488 entries. Assignments were made if a detected ion was within 3 ppm ( $\Delta m/m \times 10^6$ ) of a theoretical formula. If more than one theoretical formula met this requirement, the  $^{13}\text{C}$  containing peak intensity was considered, and the formula was chosen with the carbon number closest to that predicted by the  $^{13}\text{C}$  isotope: monoisotopic peak intensity ratio. After assignment and alignment of all samples to formulas, the formula list was cropped to only include entries that were found in at least one sample, were monoisotopic, and did not include monosulfate peaks ( $\text{SO}_3$  class), as these were found to mainly consist of contaminants. The resulting list was 5391 formulas, and each sample was represented by a matrix of these 5391 rows and columns as transient times.

Assigned intensities were summed for each transient time to obtain a chromatogram of the total assigned current (TAC; not including  $^{13}\text{C}$  peaks), and this was binned into intensity areas of 0.2 min, as for the standards, which were subsequently divided by the estimated microgram amounts of the eluting substance (estimated from the CAD bin areas) to obtain binned data of TAC/ $\mu\text{g}$ , or relative response factor.

## RESULTS AND DISCUSSION

**Relative Response Factor of Purchased Carboxylic Acids.** To compare the response of ESI-HRMS and CAD for individual analytes, a set of model compounds was tested. Because of the predominance of acidic groups in routine analysis of DOM samples (negative ESI-MS), a series of carboxylic acids

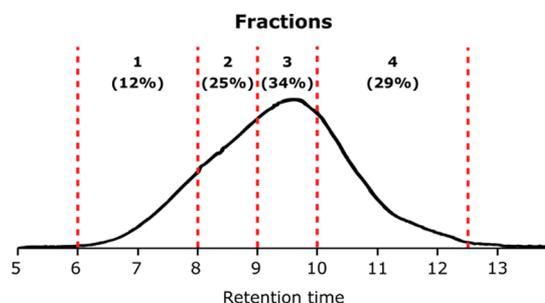
characterized by different structures, hydrophobicity (octanol–water partition coefficients;  $\log P$ ), MW, and number of carboxyl groups was investigated (Table 1). The acids eluted during the gradient mobile phase elution, and a linear relationship was obtained between the retention time ( $t_r$ ) and calculated  $\log P$  (eq 1, Figure SI3).

$$t_r = 1.96 \log P + 6.48 \quad (1)$$

As expected, the MS sensitivity was variable across the model compounds and exhibited a broad relative standard deviation among those which gave a response (RSD  $\approx$  46%), while the CAD response was characterized by a more uniform response (RSD  $\approx$  16%) (Figure 1). The CAD response had no statistically significant relationship with mass, H/C, O/C, or  $\log P$ , indicating that while slightly variable, the sensitivity was uniform for the compounds tested (Pearson's Rho  $p > 0.05$ ). Additionally, the sensitivity was found to be similar for nonacids, such as a purchased lignin mixture (Sigma-Aldrich, data not shown) and several nonacid compounds.<sup>15,20</sup> According to the central limit theorem, as mixtures become more diverse, we expect the effect of compound-to-compound variability to be minimized in the average response,<sup>21</sup> and for the CAD to become a more reliable estimator of abundance as the mixture becomes more diverse. Similarly, we posit that the MS response of a mixture is likely to tend toward the average response as the mixture becomes more diverse. MS relative response factor also had no significant trend with H/C, O/C, or  $\log P$  (Pearson's Rho  $p > 0.05$ ) but did have a very slight negative trend with mass (Pearson's Rho  $-0.57$ ,  $p < 0.05$ ); however, the mass range was highly skewed to three lower responding higher mass compounds (Table 1).

The relative response factor of the purchased acids (slope of MS response vs  $\mu\text{g}$  injected) varied depending on whether the acid was injected in a clean solvent solution or in a solution containing a high concentration of SRFA. We anticipated that the response would be suppressed, but in several cases the response was higher (Table 1), possibly due to some enhancement from altered pH or a more complex mechanism. The median response was similar ( $2.4$  vs  $2.7 \times 10^6$  counts/ $\mu\text{g}$ ), but the mean, standard deviation, and standard error of the mean increased (Table 1). The ESI-MS response (TAC) of SRFA increased with the injected amount in a near linear fashion (Figure SI4), with some curvature at high and low abundances.

**SRFA Size Fractionation and Relative Response Factor (TAC/ $\mu\text{g}$ ).** SRFA powder was dissolved and fractionated into four molecular size fractions by HPSEC (Figure 2).



**Figure 2.** SRFA profile by HPSEC–CAD. The dashed lines identify the collected fractions: fraction 1 (6–8 min); fraction 2 (8–9 min); fraction 3 (9–10 min); and fraction 4 (10–12.5 min). The percentages correspond to the amount of the total material injected expected in each fraction, based on the integral of the CAD signal.

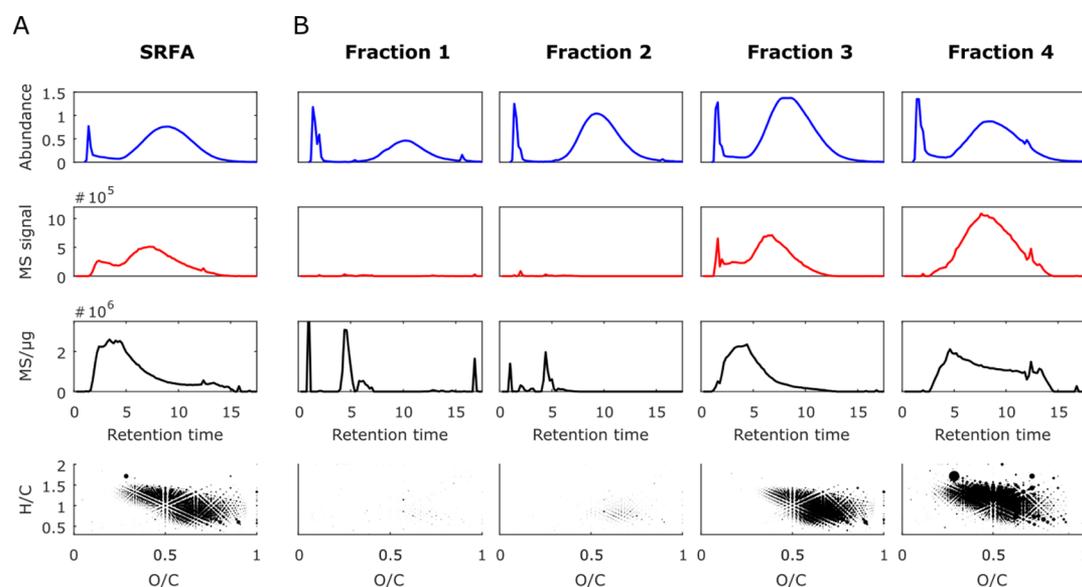
The SRFA was also prepared as a bulk solution for comparison with the size fractions. Analysis was conducted by reversed phase chromatography and detected by ESI(–)-HRMS and CAD (Figure 3). Because of the mixture's complexity, organic material was continuously eluted from the column, which eventually decreased in abundance below the detection limit of both detectors. The solvent front ( $RT < 1.8$  min), devoid of any ESI-HRMS response, was excluded from consideration. The majority of the material eluted at a retention time corresponding to  $\log P$  0–4 (Figure SI5), similar to previous results.<sup>22</sup> A clear divergence between the signals from the two detectors emerged (Figure 3A), leading to a decreasing relative response factor as more of the hydrophobic material was eluted.

The hydrophilic and mid-polarity organic material ( $RT < 10$  min;  $\log P < 1.79$ ) was responsible for the majority of the MS signal, in contrast with the more hydrophobic material ( $RT > 10$  min;  $\log P > 1.79$ ) poorly ionizable compared to the abundance observed by the CAD (Figure 3A). One explanation for this result might be that the relative response factor decreases with hydrophobicity, but this is unlikely as hydrophobic compounds are more likely to be at the surface of the droplets in the ESI spray and to enter the gas phase, and hydrophobicity has been shown to relate poorly to ionization efficiency.<sup>6</sup> More likely there is a multicomponent mixture with at least one part highly ionizable with a broad polarity range, and at least one part poorly ionizable with mainly hydrophobic species ( $\log P > 1.79$ ). The presence of material resistant or susceptible to ionization in this type of complex mixture is documented,<sup>5,23–25</sup> but the quantitative extent of these two DOM pools have not been clearly defined.

Based on previous results, we anticipated that the high MW material in SRFA would have a low ESI response.<sup>25</sup> The mobile phase blank showed a null response in both detection mode (data not shown). Fractions 1–4 represented 12, 25, 34, and 29% of the bulk SRFA, respectively (based on integration of the CAD response from a previous run; Figure 2). When re-analyzed by reversed phase HPLC-CAD-ESI-HRMS, some loss of material was observed, as only  $\sim$ 90% of fractions 1, 2, and 4 was recovered, meaning that 6% of the total material was lost during fraction treatment. The first two fractions had very little highly ionizable material, except for the presence of few low abundance  $\text{C}_x\text{H}_y\text{-SO}_3$  peaks commonly found as contaminants (e.g., plasticizers), which had a high relative response factor (Figure 3B). The third fraction had a similar relative response factor chromatogram to the bulk material, appearing as a fractal representation of the bulk and perhaps needing further separation, while the fourth fraction gave a more stable, high relative response (Figure 3B). The larger molecular size fractions (fractions 2–3) had some intensity at mass defects 0.5–0.6, in addition to the typical singly charged peaks at mass defects 0–0.3 (Figure SI6), suggesting that some higher MW doubly charged material was present, but the fourth fraction was exclusively singly charged.

Each fraction contained a material with a broad range of polarities up to a retention time of 15 min ( $\log P$  4.34). The polarity distribution of the four fractions varied slightly, with the early fractions having a higher average retention time and only the last two fractions having material eluting before 5 min ( $\log P < -0.76$ ).

The size exclusion method used to separate the fractions dictates that the material with the largest molecular size elutes first but also suffers from a charge exclusion effect that leads to early elution (relative to size) of the most charged material.<sup>25,26</sup>



**Figure 3.** (A) SRFA bulk and (B) fractions, analyzed by reverse-phase chromatography coupled to CAD (blue) and ESI-MS (red). The relative response factor (black) is expressed as assigned MS signals per  $\mu\text{g}$  of material in each time bin. The van Krevelen diagrams (bottom panels) show the molecular atomic ratios (H/C vs O/C) of each assigned formula in the highly ionizable material, scaled to signal intensity.

This probably explains why some low MW but highly charged (several highly acidic carboxylic acids) compounds elute in fraction 3 instead of fraction 4. In the van Krevelen diagrams, it is apparent that the ionizable material in fraction 3 is the most oxygenated and acidic material, located at the bottom right corner of the diagram.<sup>27</sup> Overall, these results suggest that the highly ionizable component of complex DOM mixtures like SRFA is low MW and relatively hydrophilic, while the poorly ionizable fraction has a high MW and on average more hydrophobic. In the unfractionated sample, the two pools (highly and poorly ionizable) elute together in reversed phase HPLC but are slightly offset, leading to the decreasing relative response factor observed in Figure 3A.

**Model to Estimate the Abundance of Highly and Poorly Ionizable DOM.** SRFA fractionation demonstrated the presence of two broad pools; highly and poorly ionizable materials. At least 33% of the bulk SRFA material was poorly ionizable by ESI, according to the sum of the material recovered in fractions 1 and 2. The results proved that at least a third of the SRFA material was resistant to negative-mode ionization (null response factor), while the highly ionizable component, not completely isolated with the fractionation, was distributed between the last two fractions and the nonrecovered material.

The relative response factor, namely, the ability of the material to ionize in the set conditions, could be used as the parameter to further discriminate between highly and poorly ionizable components in a mixture. We estimated the abundance of the highly ionizable material in a sample by assuming that the average negative-mode ESI response of a mixture of unknown acids (the highly ionizable material) would be the same as the average response of the acids we purchased (i.e.,  $2.6 \times 10^6$  MS counts  $\times$  time/ $\mu\text{g}$  material). This is a clear simplification based on the central limit theorem, stating that the average of a large number of independent variables with a normal distribution gives a result approaching the arithmetic mean result.<sup>21,28</sup> The important assumptions of this approach are that the chosen acids are representative of acids in natural aquatic mixtures, both have normal distributions of the relative response factor, and ionization suppression and promotion balance out so as not to

affect the average result. The model also assumes that the response is linear with concentration, but lower concentration in the ESI spray may lead to a higher relative response because of less competition for the charge. Indeed, there was a slight curvature at low and high injected abundances for our SRFA calibration (Figure S14), but the response was almost linear over a wide range that covers most of the bin abundances in this study. Interestingly, the effect of this curvature means that the samples and fractions that had a low relative response were actually overestimated, and those with high response were underestimated, so our results are conservative in terms of differences between samples and fractions.

With our experiments, we cannot critically assess these assumptions for a complex mixture like SRFA but take this as a first approach to the estimation of the abundance of highly ionizable material in the mixture. The two-component mixed model is summarized in eq 2

$$\text{RF}_S \times m_S = \text{RF}_H \times m_H + \text{RF}_P \times m_P \quad (2)$$

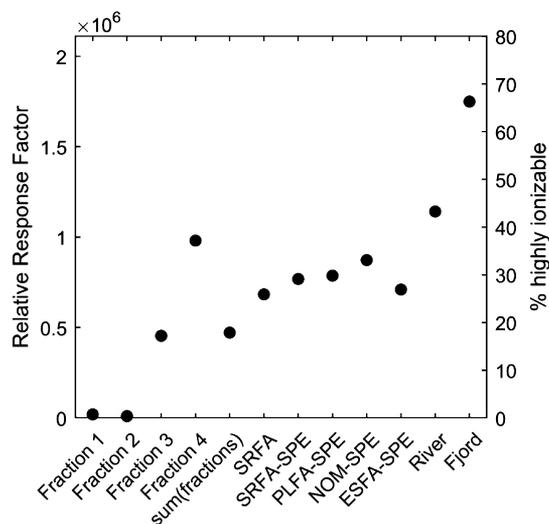
where RF is the relative response factor and the subscripts define the different components: sample time bin (S), highly ionizable (H), and poorly ionizable (P);  $m$  signifies the amount of the material in the time bin and  $\text{RF}_S$  represents the relative response factor of the material in the time bin. For simplicity, the highly ionizable material ( $m_H$ ) was expressed as a percentage; therefore,  $m_S$  was set to 100; additionally, the second term of eq 2 is simplified because  $\text{RF}_P$  is zero, leading to the eq 3

$$m_H (\%) = (\text{RF}_S / \text{RF}_H) \times 100 \quad (3)$$

Equation 3 was used to calculate the proportion of the highly ionizable material in each time bin of the chromatogram. The full relative response factor for each sample ( $\text{RRF}_{\text{sample}}$ ) was calculated by calculating a weighted average of the response in each time bin according to eq 4, where a value of 100% would signify the sample was entirely composed of ionizable acids.

$$\text{RRF}_{\text{sample}} (\%) = \frac{\sum (\text{RF}_S \times m_S)}{\text{RF}_H \sum m_S} \times 100 \quad (4)$$

The model was applied to estimate  $RRF_{\text{sample}}$  for the four fractions and to the bulk material (Figure 4). Fractions 1 and 2

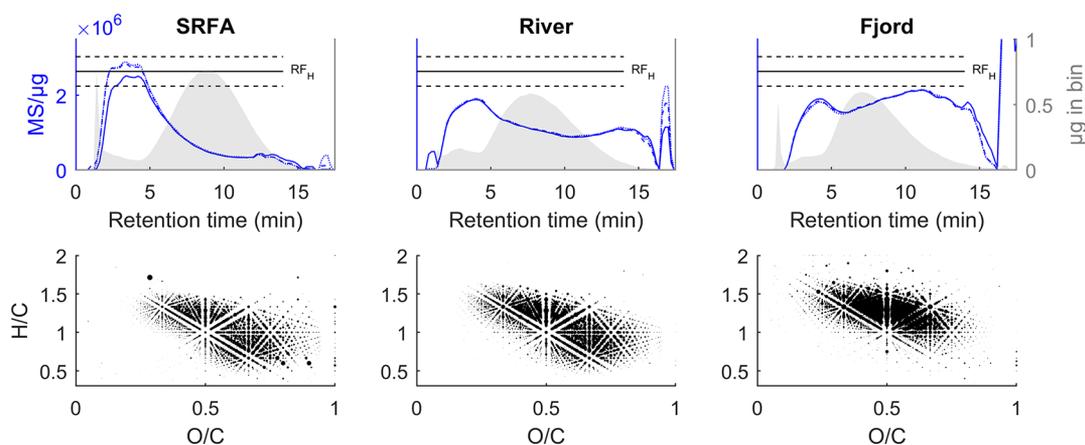


**Figure 4.** Application of the two component mixed model to the samples (black dots). Relative response factor (TAC/ $\mu\text{g}$ , primary y-axis); highly ionizable material (percentage, secondary y-axis). The sample points apply to both y-axes, and in each case represent a single measurement, so error bars are not available. Triplicate analysis of SRFA-SPE, river and fjord, suggested that the standard deviation was less than  $0.014 \times 10^6$  RRF, so smaller than the plotted points. Sample codes: Suwannee River Fulvic Acid (SRFA), Pony Lake Fulvic Acid (PLFA), Nordic Natural Organic Matter (NOM), Elliot Soil Fulvic Acid (ESFA).

were confirmed to be characterized exclusively by the poorly ionizable material (except a negligible percentage in fraction 1 lower than 1%). The model revealed that both fractions 3 and 4 had a mixed nature, respectively, 17 and 41% of the eluting material was recognized as highly ionizable. These results were combined to estimate the amount of ionizable DOM in SRFA, corresponding to just 17% of the total unfractionated material, in contrast with the 77% of poorly ionizable material across the four fractions. The final 6% was constituted by the material lost

during handling and analysis of samples, whose ESI response is unknown. These results suggest that an impressive amount of material in the SRFA mixture is uncharacterized in routine ESI-MS analysis performed in negative mode. It appears that the totality of the high MW material (fraction 1 and 2, Figure 4) and a large part of the molecules in the typical DOM molecular size range (fractions 3 and 4) is resistant to negative ionization. Obviously, the model is only an approximation, a clear distinction between the highly and poorly ionizable material would only be possible if a complete isolation between the two components was achieved, and the effect of ionization suppression was completely removed or understood. Nevertheless, these results highlight the importance of further investigation on the nature of the material elusive to negative ESI analysis, and a clear consideration in HRMS studies about the extent of sample coverage that the ionization technique allows.

The model was also applied to the bulk SRFA (before and after SPE on Agilent PPL) on the basis of its averaged relative response factor. The results revealed a larger abundance of highly ionizable species in the unfractionated SRFA compared to the combined fractions (Figure 4). In fact, the percentage of highly ionizable DOM in the unfractionated SRFA was 27% (and 30% in the SPE mixture), in clear contrast with the estimation of the highly ionizable material from the analysis of the isolated fractions (17%). In order to explore the differences between the combined fractions and the full SRFA (before and after SPE), the evolution of their relative response factors along the elution time was compared (Figure S17). Only a poor match was obtained among the profiles at low retention times ( $RT < 6$  min) where about 8–9% of the total injected material eluted. This suggests that the material lost during the preparation of the four fractions was hydrophilic and had a high relative response factor. The differences between the combined fractions and the full SRFA emphasize the fact that ESI of complex mixtures is complicated. Even considering the amount of material lost during the sample manipulations (recovery of the fractions), this discrepancy could not be explained. Possibly, synergy among the different species in the unfractionated material was responsible for the increase in the relative response factor, due to a complex ionization promotion, as observed for some of the purchased



**Figure 5.** Solid-phase extracted samples analyzed by reverse-phase chromatography coupled to ESI-MS and CAD: SRFA (Left), river (Middle), and fjord (Right). The blue profiles show the relative response factor (TAC/ $\mu\text{g}$ ) variation as the function of the elution time (three replicates overlaid), the shaded grey areas show the material abundance from CAD. A horizontal black line indicates the constant  $RF_H$  of the purchased acids, the horizontal dotted black lines are standard error of the mean. The van Krevelen diagrams (bottom panels) show the molecular distribution of the highly ionizable material. Point size indicates signal intensity.

acids when diluted into a complex mixture (Table 1). These effects resulted in the larger estimation of the highly ionizable material in the bulk sample compared to combined fractions from the model.

It has previously been demonstrated that Agilent PPL is able to retain the majority of ionizable DOM in aquatic samples,<sup>29</sup> but interestingly, the SPE cartridge also retained the majority of the poorly ionizable material. We had supposed that this larger, polymeric material might be trapped in the frit of the cartridge or be irreversibly retained, but found only a slight increase in the relative response factor at later retention times after SPE (Figure S18). The fact that Agilent PPL retains both highly and poorly ionizable materials suggests that the typically observed SPE extraction efficiency (approximately 60%<sup>30,31</sup>) contains both these pools, and the amount of highly ionizable DOM in aquatic samples is somewhat less than 60%. We also measured three other reference materials provided by the International Humic Substances Society (PLFA, ESFA and Nordic Reservoir NOM) and found similar results (Figures 4 and S18). Note that we do not assess the ionization efficiency of material that is not extracted by PPL or other preparation resins used by the International Humic Substances Society, but previous reports suggest that the non-extracted material is poorly ionizable.<sup>29</sup>

**DOM Relative Response Factor in Terrestrial and Marine Settings.** Solid-phase extracted DOM samples of terrestrial origins (SRFA and a river) were compared to a marine sample taken from a Swedish fjord. The relative response factor profiles of the samples from the three locations changed considerably from land to sea (Figures 5 and S19). The SRFA sample (representative of a swamp end-member) was the most affected by hydrophobic, poorly ionizable material, and the high-order river (Figure 5) more closely matched Fraction 4 of SRFA (Figure 3). This suggests a moderate contribution from hydrophobic, poorly ionizable DOM in the river, but less so compared with the material in SRFA, which has a more “humified” terrestrial material not previously exposed to as many aquatic degradation processes. The fjord showed a stable and higher relative response factor, almost constant throughout the elution range (Figure 5); differently from the terrestrial samples, an increase in the relative response factor profile was observed for the more retained material ( $RT > 8$  min;  $\log P > 0.77$ ), suggesting the predominance of highly ionizable DOM throughout the polarity gradient. The average response factor value of the fjord-DOM was also the closest to the average of the carboxylic acids (Figure 5), suggesting an abundance of small acidic molecules.

Application of the two component model to these solid phase extract samples suggested that 69% of the fjord-SPEDOM was highly ionizable, in strong contrast with the 44 and 30% of the highly ionizable material in the river-SPEDOM and in SRFA, respectively (Figure 4). Similar to SRFA, other reference mixtures (highly representative of terrestrial-DOM) showed a limited abundance of the highly ionizable material (Figure S18), indicating that terrestrial DOM is strongly affected by the presence of a poorly ionizable component. These results seem to confirm a progressive loss of the poorly ionizable component from land to sea. Taken together with previous literature on this topic, we suggest that our findings indicate that relatively hydrophobic carboxylic-rich alicyclic-type moieties increase in relative abundance as DOM is degraded in water bodies, while lignin polymers, which attenuate the most light and are higher MW, are gradually removed.<sup>25,32–34</sup>

**Further Considerations.** Three main caveats are necessary to understand our interpretations of the data presented:

- 1 The model assumes that the tested carboxylic acids (Tables 1 and S11) were representative for the response of the ionizable acidic molecules in the complex mixture. This assumption is obviously debatable. Some acids (such as ibuprofen) can generate no signal when analyzed by ESI(–)-HRMS and conversely, noncarboxylic acid compounds (such as fraxin<sup>3</sup>), produce a strong response. However, the behavior of single molecules is not representative of a complex mixture like DOM, where the resulting signal derives from the mixture average according to the central limit theorem.<sup>21,28</sup> Further studies that test and model a larger number of compounds would be useful. Unfortunately, molecules that are highly representative of organic molecules rich in acidic functionalities and analogous to the expected DOM composition (i.e., carboxyl rich alicyclic molecules) are not currently available for purchase.
- 2 In the model, we assume that different complex mixtures of acidic compounds will have the same averaged relative response factor. However, the fjord and river samples have differing levels of saturation (H/C ratio of the detected compounds; Figure 5), and if this feature has a broad effect on the response factor, this would change the interpretation of the model in Figures 4 and 5. We found no significant relationship between saturation and response for the carboxylic acid standards (Pearson's Rho  $p > 0.05$ ), but this possibility deserves further attention.
- 3 The effects of ionization suppression and promotion are complex. We determined that some purchased carboxylic acids were suppressed and some were promoted when diluted into SRFA, leaving the average response fairly similar (Table 1)—however, further work investigating the ionization suppression effects of dilute compounds within DOM mixtures would be valuable.

## CONCLUSIONS

The ionization and analysis of complex mixtures are not trivial. The sample variability and the bias affecting the ionization mechanisms preclude the characterization of part of the material. Our results suggest that the extent of this bias can be numerically estimated.

The results of fractionation and re-analysis of SRFA suggested that at least the highest MW third of the material in the mixture was constituted by poorly ionizable species and the remaining lower MW material was a mixture of highly and poorly ionizable components. A model was proposed in order to discriminate between the material elusive and susceptible to ionization, and the model revised our estimate to about two-thirds of bulk-SRFA being unresponsive to negative ESI. The model suggested also that other terrestrially derived reference mixtures were afflicted by the same prominence of the poorly ionizable material, persistently present in the samples even following SPE. Increased separation of complex mixtures (e.g., moving from direct infusion to chromatography, or otherwise fractionating the sample) tends to increase the apparent proportion of material that is poorly ionizable, suggesting a valuable gain of information by sample fractionation.

Additionally to the reference mixtures, two samples isolated from a high order river and a fjord were analyzed. A clear trend emerged, suggesting that the processing of DOM along the aquatic continuum from peat systems to the sea leads to an increasing portion of the highly ionizable material in the sample, suggesting a preferential loss of poorly ionizable (high MW) lignin material as dissolved organic carbon concentration decreases. This scenario could have important repercussions on the study of the biogeochemical processes and water treatment, and further investigations about this abundant material elusive to negative ESI are indispensable.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.analchem.0c03438>.

CAD calibration curves, ESI(−)-MS calibration curves, relationship between calculated log *P* and retention time of carboxylic acids, calibration of ESI-MS response, distribution of material in SRFA, comparison among signals at 0.5–0.6 Da mass defect isolated from fractions 2, 3, and 4, evolution of the response factor, CAD, ESI(−)-HRMS, relative response factor profiles, and van Krevelen diagrams (PDF)

ESI and ion optic tune settings (XLSX)

Purchased carboxylic acid information (XLSX)

## ■ AUTHOR INFORMATION

### Corresponding Author

Jeffrey A. Hawkes – Analytical Chemistry, Department of Chemistry—BMC, Uppsala University, 75237 Uppsala, Sweden; [orcid.org/0000-0003-0664-2242](https://orcid.org/0000-0003-0664-2242); Email: [jeffrey.hawkes@kemi.uu.se](mailto:jeffrey.hawkes@kemi.uu.se)

### Authors

Claudia Patriarca – Analytical Chemistry, Department of Chemistry—BMC, Uppsala University, 75237 Uppsala, Sweden; [orcid.org/0000-0002-4301-3923](https://orcid.org/0000-0002-4301-3923)

Andrea Balderrama – Analytical Chemistry, Department of Chemistry—BMC and Limnology, Department of Ecology and Genetics, Uppsala University, 75237 Uppsala, Sweden

Martina Moze – Analytical Chemistry, Department of Chemistry—BMC, Uppsala University, 75237 Uppsala, Sweden

Per J. R. Sjöberg – Analytical Chemistry, Department of Chemistry—BMC, Uppsala University, 75237 Uppsala, Sweden

Jonas Bergquist – Analytical Chemistry, Department of Chemistry—BMC, Uppsala University, 75237 Uppsala, Sweden

Lars J. Tranvik – Limnology, Department of Ecology and Genetics, Uppsala University, 75236 Uppsala, Sweden

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.analchem.0c03438>

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

C.P., J.B., and L.J.T. were funded by the Knut and Alice Wallenberg Foundation (grant 2013.0091), J.A.H. was funded by the Swedish Research Council (VR Grant 2018-04618), A.B. is grateful for the Swedish Institute Study Scholarship (11824/2017) funded by the Swedish Institute. The authors are grateful

to two anonymous reviewers whose insights improved the manuscript.

## ■ REFERENCES

- (1) Hertkorn, N.; Frommberger, M.; Witt, M.; Koch, B. P.; Schmittkopplin, P.; Perdue, E. M. *Anal. Chem.* **2008**, *80*, 8908–8919.
- (2) Hawkes, J. A.; Patriarca, C.; Sjöberg, P. J. R.; Tranvik, L. J.; Bergquist, J. *Limnol. Oceanogr. Lett.* **2018**, *3*, 21–30.
- (3) Patriarca, C.; Bergquist, J.; Sjöberg, P. J. R.; Tranvik, L.; Hawkes, J. A. *Environ. Sci. Technol.* **2018**, *52*, 2091–2099.
- (4) Spranger, T.; Pinxteren, D. v.; Reemtsma, T.; Lechtenfeld, O. J.; Herrmann, H. *Environ. Sci. Technol.* **2019**, *53*, 11353–11363.
- (5) These, A.; Reemtsma, T. *Anal. Chem.* **2003**, *75*, 6275–6281.
- (6) Krueve, A.; Kaupmees, K.; Liigand, J.; Leito, I. *Anal. Chem.* **2014**, *86*, 4822–4830.
- (7) Witt, M.; Fuchser, J.; Koch, B. P. *Anal. Chem.* **2009**, *81*, 2688–2694.
- (8) Perdue, E. M.; Ritchie, J. D. *Dissolved Organic Matter in Freshwaters. Treatise on Geochemistry*, 2nd ed.; Elsevier, 2013.
- (9) Gross, J. H. *Mass Spectrometry-A Textbook*, 2nd ed.; Springer Heidelberg; 2011.
- (10) Reuter, H.; Gensel, J.; Elvert, M.; Zak, D. *Anal. Chem.* **2017**, *89*, 13449.
- (11) Villacorte, L. O.; Ekowati, Y.; Neu, T. R.; Kleijn, J. M.; Winters, H.; Amy, G.; Schippers, J. C.; Kennedy, M. D. *Water Res.* **2015**, *73*, 216–230.
- (12) Huber, S. A.; Balz, A.; Abert, M.; Pronk, W. *Water Res.* **2011**, *45*, 879–885.
- (13) Duarte, R. M. B. O.; Barros, A. C.; Duarte, A. C. *J. Chromatogr. A* **2012**, *1249*, 138–146.
- (14) Rojas, A.; Sandron, S.; Wilson, R.; Davies, N. W.; Haddad, P. R.; Shellie, R. A.; Nesterenko, P. N.; Paull, B. *Anal. Chim. Acta* **2016**, *909*, 129–138.
- (15) Vehovec, T.; Obreza, A. *J. Chromatogr. A* **2010**, *1217*, 1549–1556.
- (16) Gamache, P. H. *Charged Aerosol Detection for Liquid Chromatography and Related Separation Techniques*; John Wiley & Sons, 2017.
- (17) Gamache, P. H.; McCarthy, R. S.; Freeto, S. M.; Asa, D. J.; Woodcock, M. J.; Laws, K.; Cole, R. O. *LCGC North Am.* **2005**, *23*, 101–106.
- (18) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. *Limnol. Oceanogr.: Methods* **2008**, *6*, 230–235.
- (19) Górecki, T.; Lynen, F.; Szucs, R.; Sandra, P. *Anal. Chem.* **2006**, *78*, 3186–3192.
- (20) Ligor, M.; Studzińska, S.; Horna, A.; Buszewski, B. *Crit. Rev. Anal. Chem.* **2013**, *43*, 64–78.
- (21) Zark, M.; Christoffers, J.; Dittmar, T. *Mar. Chem.* **2017**, *191*, 9–15.
- (22) Namjesnik-Dejanovic, K.; Cabaniss, S. E. *Environ. Sci. Technol.* **2004**, *38*, 1108–1114.
- (23) Putman, J. C.; Gutiérrez Sama, S.; Barrère-Mangote, C.; Rodgers, R. P.; Lobinski, R.; Marshall, A. G.; Bouyssière, B.; Giusti, P. *Energy Fuels* **2018**, *32*, 12198–12204.
- (24) Rodgers, R. P.; Mapolelo, M. M.; Robbins, W. K.; Chacón-Patino, M. L.; Putman, J. C.; Niles, S. F.; Rowland, S. M.; Marshall, A. G. *Faraday Discuss.* **2019**, *218*, 29.
- (25) Hawkes, J. A.; Sjöberg, P. J. R.; Bergquist, J.; Tranvik, L. J. *Faraday Discuss.* **2019**, *218*, 52–71.
- (26) Mori, S. *Anal. Chem.* **1989**, *61*, 530–534.
- (27) Zhrebek, A.; Kostyukevich, Y.; Kononikhin, A.; Kharybin, O.; Konstantinov, A. I.; Zaitsev, K. V.; Nikolaev, E.; Perminova, I. V. *Anal. Bioanal. Chem.* **2017**, *409*, 2477–2488.
- (28) Miller, J. N.; Miller, J. C. *Statistics and Chemometrics for Analytical Chemistry*; Harlow, 2010.
- (29) Raeke, J.; Lechtenfeld, O. J.; Wagner, M.; Herzsprung, P.; Reemtsma, T. *Environ. Sci.: Processes Impacts* **2016**, *18*, 918–927.

(30) Green, N. W.; Perdue, E. M.; Aiken, G. R.; Butler, K. D.; Chen, H.; Dittmar, T.; Niggemann, J.; Stubbins, A. *Mar. Chem.* **2014**, *161*, 14–19.

(31) Li, Y.; Harir, M.; Lucio, M.; Kanawati, B.; Smirnov, K.; Flerus, R.; Koch, B. P.; Schmitt-Kopplin, P.; Hertkorn, N. *Anal. Chem.* **2016**, *88*, 6680–6688.

(32) Massicotte, P.; Asmala, E.; Stedmon, C.; Markager, S. *Sci. Total Environ.* **2017**, *609*, 180–191.

(33) Köhler, S. J.; Kothawala, D.; Futter, M. N.; Liungman, O.; Tranvik, L. *PLoS One* **2013**, *8*, No. e70598.

(34) Hertkorn, N.; Benner, R.; Frommberger, M.; Schmitt-Kopplin, P.; Witt, M.; Kaiser, K.; Kettrup, A.; Hedges, J. I. *Geochim. Cosmochim. Acta* **2006**, *70*, 2990–3010.