

LETTER

Extreme isomeric complexity of dissolved organic matter found across aquatic environments

Jeffrey A. Hawkes ^{1*}, Claudia Patriarca,¹ Per J. R. Sjöberg,¹ Lars J. Tranvik ², Jonas Bergquist¹

¹Analytical Chemistry, Department of Chemistry - BMC, Uppsala University, Uppsala, Sweden; ²Department of Limnology, Uppsala University, Uppsala, Sweden

Scientific Significance Statement

Molecular complexity is an inherent feature of dissolved organic matter (DOM), confounding investigations into its nature. Recent research has suggested that this complexity may explain the persistence of DOM due to the implied low abundance of individual compounds. Here, we use chromatographic separation and collision induced dissociation of deprotonated molecules to demonstrate the extreme isomeric complexity of individual molecular formulas in DOM and show that isomeric complexity occurs across diverse aquatic environments.

Abstract

The natural aquatic environment contains an enormous pool of dissolved reduced carbon, present as ultra-complex mixtures that are constituted by an unknown number of compounds at vanishingly small concentrations. We attempted to separate individual structural isomers from several samples using online reversed-phase chromatography with selected ion monitoring/tandem mass spectrometry, but found that isomeric complexity still presented a boundary to investigation even after chromatographic simplification of the samples. However, it was possible to determine that the structural complexity differed among samples. Our results also suggest that extreme structural complexity was a ubiquitous feature of dissolved organic matter (DOM) in all aquatic systems, meaning that this diversity may play similar roles for recalcitrance and degradation of DOM in all tested environments.

Dissolved organic matter (DOM) is the dominant form of organic carbon in most aquatic environments. Upon mineralization, it is an important precursor of outgassing of CO₂

from inland waters (Tranvik et al. 2009), it carries substantial amounts of nutrients and energy from land to sea (Medeiros et al. 2016), and it persists in the deep ocean for millennia (Dittmar and Stubbins 2014). It is an ultra-complex mixture of compounds that provides the ultimate test of the capabilities of analytical chemistry (Rodgers et al. 2005; Dittmar and Stubbins 2014), and investigations into its nature have been confounded by its extreme molecular complexity. High-resolution mass spectrometry (HRMS) is able to resolve many thousands of molecular masses from complex natural mixtures of organic compounds (Marshall et al. 1998; Riedel and Dittmar 2014; Hendrickson et al. 2015), but is unable to differentiate between structural isomers of a molecular formula.

Most recent research has utilized advanced visualization or multivariate statistical approaches to interpret HRMS data (Wu et al. 2004; Sleighter et al. 2010; Kellerman et al. 2015), but when the HRMS data are not combined with more

*Correspondence: jeffrey.hawkes@kemi.uu.se

Author Contribution Statement: JAH designed the research and wrote the MATLAB code. JAH, CP, and PJS conducted the experiments. LJT and JB acquired equipment and funding. JAH wrote the paper with contributions from all authors. JAH is accountable for the integrity of the data.

Data Availability Statement: Data are available in the repository at <https://doi.pangaea.de/10.1594/PANGAEA.885274>.

Additional Supporting Information may be found in the online version of this article.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

structurally sensitive techniques (Lam et al. 2007; Arakawa and Aluwihare 2015; Arakawa et al. 2017), average biogeochemical behavior of several structural isomers is measured without knowledge of how diverse the isomeric mixture actually is (Zark et al. 2017). The study of organic matter in the environment relies on detecting and characterizing subtle changes to the complex mixture, but most of the complexity of the mixture is hidden behind this isomeric averaging.

The isomeric complexity within each molecular formula may be probed by chromatography or by fragmentation within the mass spectrometer. These techniques have previously been used separately to confirm that numerous structural isomers contribute to any obtained molecular formula (Leenheer et al. 2001; Witt et al. 2009; Capley et al. 2010; Arakawa and Aluwihare 2015; Cortés-Francisco and Caixach 2015; Brown et al. 2016; Zark et al. 2017). Chromatographic separation of complex mixtures of DOM has revealed variability in the primary and tertiary structure of their components (Saleh et al. 1989; Woelki et al. 1997; Reemtsma and These 2003; Namjesnik-Dejanovic and Cabaniss 2004; Woods et al. 2011; Arakawa et al. 2017), and several recent studies have led to tremendous advances in our understanding of the compositional complexity of terrestrial humic and fulvic acids via chromatographic separation and subsequent high resolution MS analysis (Koch et al. 2008; Stenson 2008; Gaspar et al. 2010; Brown et al. 2016; Sandron et al. 2017). Fragmentation has revealed the importance of carboxylic acid and alcohol functionality in fulvic acids (Leenheer et al. 2001; Witt et al. 2009; Capley et al. 2010; Zark et al. 2017) and has led to estimates that at least 28 structural isomers are present at any particular molecular mass in seawater (Zark et al. 2017). When used in combination with chromatography, fragmentation has allowed the separation and measurement of specific biomarkers, metabolites, or contaminants within a complex mixture (Wu et al. 2015; Longnecker and Kujawinski 2017), but has rarely been used to probe the isomeric complexity of the bulk organic mixture (Brown et al. 2016).

Online coupling of high performance liquid chromatography (HPLC) with HRMS or HRMS/MS have been underexploited in the investigation of natural ultra-complex mixtures (Patriarca et al. 2017; Petras et al. 2017). This is partly due to the nature of electrospray ionization, the most commonly applied ionization method coupled to HRMS, which is sensitive to inorganic buffers and requires close control of the solvent mixture for stable ionization efficiency, meaning that typical highly buffered gradients of solvents are inappropriate (Koch et al. 2008). Also, signal to noise is typically quite low in individual transients (Brown et al. 2016), making the scan rate a critical limiting feature of online HPLC-HRMS. Collection of fractions and pre-concentration is a sensible, though time consuming, way of improving sensitivity (Capley et al. 2010; Brown et al. 2016). However, subtle detail in the elution gradient is lost through combining eluted material into fractions (Brown et al. 2016).

Here, we use an online HPLC-tandem HRMS method to explore the isomeric complexity of individual molecular masses in several natural samples. Previous studies have demonstrated that similar structural complexity is found at every tested molecular mass in such complex mixtures (Witt et al. 2009; Capley et al. 2010; Zark et al. 2017), so here we do not explore numerous masses, leaving this for future investigation according to individual research questions. Instead, we go into greater detail at one particular nominal mass, $[M-H]^- = 369$, which is near the average molecular mass in most aquatic systems, and for which three soluble and commercially available compounds could be purchased and compared to the natural mixtures (Fig. 1). We test the method on extracted organic matter from several aquatic environments, ranging from a headwater stream to the deep ocean.

Experimental

Reagents and instrumentation

All solvents were high purity grade (Supporting Information) and glassware was muffled at 450°C for at least 4 h prior to use. The HPLC was an Agilent 1100 with binary pump and autosampler. The Orbitrap was an LTQ-Velos-Pro (Thermo Scientific, Germany). Model compounds A and B (Fig. 1) were purchased from Sigma Aldrich (Sweden) and compound C was purchased from Toronto Research Chemicals (Canada), as powders. Suwannee River fulvic acid (SRFA) and Nordic Reservoir natural organic matter (NRNOM) reference materials were purchased from the International Humic Substances Society (U.S.A.).

Sample preparation

Samples were collected from a headwater stream (HW), a river (RIV) and a fjord (FJO) in Sweden, the deep Caribbean Sea (MAR), and from two terrestrial reference materials (SRFA and NRNOM). DOM from the samples was concentrated to 500 ppm C in 0.1% formic acid by solid phase extraction. The model compounds A–C were added at 10 ppb to the SRFA sample. Details about sample collection and preparation can be found in the Supporting Information.

Chromatographic methods

Separation was conducted with an Agilent PLRP-S poly(styrene/divinylbenzene) column, similar in chemistry to Agilent PPL sorbent, with dimensions 1.0×150 mm, $3 \mu\text{m}$ bead size, 100 \AA pore size, with a pre-column filter ($0.5 \mu\text{m}$, Supelco ColumnSaver). The mobile phases were 0.1% formic acid in water (mobile phase A) and 0.1% formic acid in acetonitrile (mobile phase B), in a stepped gradient from 5% to 90% B over 20 min at $50 \mu\text{L min}^{-1}$ as shown in Supporting Information Table S1.

Orbitrap-MS/MS analysis

The Orbitrap LTQ-Velos was calibrated and tuned to maximize the peak at 369.1 in SRFA (see Supporting Information). Ions were filtered to a 1 Da mass window

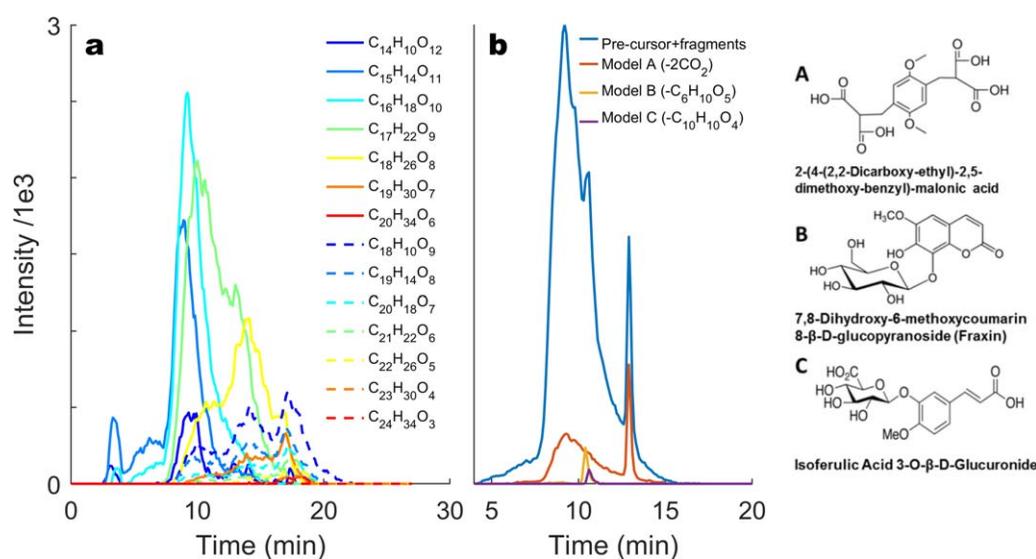


Fig. 1. (a) Chromatographic spreading of the 14 isomeric mixture peaks in the NRNOM sample (500 ppm C), as shown by summed intensities of the precursors and fragments at 21% normalized intensity in the CID cell. (b) Summed intensity of $C_{16}H_{18}O_{10}$ (third peak from a) and assigned fragments in SRFA (500 ppm C) at 21% normalized energy in the CID cell, with the model compounds added (10 ppb C each). Fragments that are specific to model compounds B and C are shown. No unique fragment was found for compound A, so a common fragment that is particularly abundant for compound A is shown. This highlights the full chromatographic separation of compound A from B and C, which overlap with each other. In contrast, the natural isomeric mixture has no easily identified features. A–C: Model compounds with formula $C_{16}H_{18}O_{10}$ and negative ion mass 369.08272.

(368.5–369.5) in a dual pressure ion trap, and fragmentation was conducted by collision induced dissociation (CID; activation q 0.25, time 10 ms) with nitrogen gas at excitation voltages of 0%, 21%, and 27% normalized collision energy. The resulting mixture of ions was transferred to the Orbitrap for analysis at resolution setting 100,000, with an accumulation time set to a maximum of 2 s and a maximum of 5×10^4 ions.

Fourteen precursor peaks with nominal m/z 369 were identified in the analysis with no excitation voltage. The exact mass of possible fragments from these precursors with neutral losses of up to five CO_2 and/or H_2O molecules were calculated to make a target mass list for assignments in the analyses with excitation voltage applied. Assignments were allowed with mass error < 3 ppm (see Supporting Information for more detail on assignment). These losses (CO_2 , H_2O) are the only ones that can be unambiguously assigned to a precursor peak due to the high likelihood of their loss and previous experiments demonstrating their dominance from single isolated ions (Witt et al. 2009). These losses made up more the 70% of total intensity (Fig. 2).

Results and discussion

Separation of precursor peaks by liquid chromatography

The 14 precursor ions with masses ranging from 369.0099 to 369.2435 and their main fragments were easily mass-resolved by the Orbitrap under the analytical conditions, but

were not fully chromatographically resolved by our HPLC method. Rather, they were partially separated with substantial overlap (Fig. 1a). The formulas with higher oxygen abundance in each series eluted first, as is expected due to their higher oxygen functionality (e.g., carboxylic acid and alcohol groups) and resulting higher average polarity (Saleh et al. 1989; Namjesnik-Dejanovic and Cabaniss 2004). Three purchased compounds with the molecular formula $C_{16}H_{18}O_{10}$ (deprotonated mass 369.0827) and functionality that resembles moieties found in natural organic matter (Hertkorn et al. 2006; Lam and Simpson 2009; Woods et al. 2011) were obtained to compare with the compounds with the same formula in the natural mixture of SRFA. In contrast to the broad humps found eluting from the complex natural mixture, the model compounds gave relatively sharp, well-resolved peaks (Fig. 1b). This showcases the extreme isomeric complexity of the natural mixture (Stenson 2008; Capley et al. 2010; Arakawa et al. 2017; Sandron et al. 2017), revealing that each precursor peak has a continuum of structural isomers that can be somewhat smeared out on such a hydrophobicity/acidity gradient. As a result, some isomers of one isomeric mixture may elute later than isomers of the next-eluting isomeric mixture. It is unclear from these results how many isomers exist per molecular formula, but our results support the estimate of some large number > 30 computed by Zark et al. (2017), assuming that individual isomers elute as well resolved chromatographic peaks, as do the model compounds (Fig. 1b and Supporting Information Fig. S3).

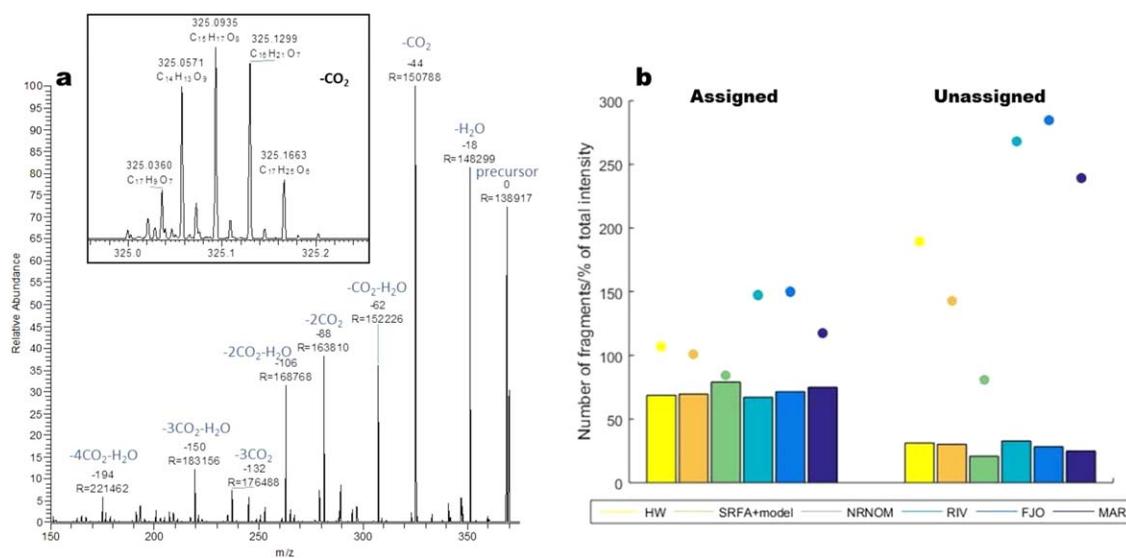


Fig. 2. (a) Integrated mass spectrum of sample HW after fragmentation at 27% normalized energy in CID mode from 2 min to 25 min, shown is the mass of the peak relative to 369 and the resolution (R) of the peak. The complete ion packet at $m/z = 369$ is included, and the main neutral losses are multiples of CO_2 (44) and H_2O (18). Inset: close up of mass 325 (precursor-44). (b) The number of fragments assigned and unassigned per sample (points) and their % contribution to the total intensity (bars). Typically, there were more fragment peaks unassigned than assigned, but the unassigned peaks made up < 30% of total intensity.

Collection of fractions followed by secondary, orthogonal separation on a second column may allow further separation of these isomers (Arakawa and Aluwihare 2015; Brown et al. 2016; Arakawa et al. 2017; Sandron et al. 2017), but this is beyond the scope of this article. Here, we consider the fragmentation patterns of each isobaric peak at $[\text{M}-\text{H}]^- 369$ in this faster (34 min), more automated method.

Fragmentation of precursor ions

The fragmentation patterns (the relative intensities of the various fragments) were dominated by neutral losses of CO_2 (-44) and H_2O (-18), as has been observed previously (Leenheer et al. 2001; Plancque et al. 2001; Witt et al. 2009; Brown et al. 2016; Zark et al. 2017) (Fig. 2). Samples that had been stored for several years at -20°C in methanol also gave substantial neutral losses of methanol (-32), indicative of methylation of carboxylic acid groups. Such samples were not further used in this study.

Model compounds B and C had glycosidic bonds, leading to fragmentation at the glycoside linkage (Fig. 1). This type of fragmentation was not observed in high abundance in the natural aquatic samples using CID (Fig. 2), suggesting that this functionality is not a major component of the natural material (Witt et al. 2009). Solid phase extraction at pH 2 is selective toward humic substances over sugars, so this may reflect a bias in our sample treatment—but we have found that glycoside compounds retain well on reverse phase sorbents at low pH, provided they have some acidic or phenolic functionality. Due to the high lability of carbohydrates, it seems likely that natural mixtures are dominated by

aromatic or aliphatic hydrocarbon backbone molecules substituted with carboxylic acid and alcohol groups that are more recalcitrant to biotic degradation (Lam and Simpson 2009; Witt et al. 2009; Dittmar and Stubbins 2014; Arakawa and Aluwihare 2015; Arakawa et al. 2017), more like the model compound A. This has been suggested previously by (Witt et al. 2009) for the same model compound, and somewhat supports the concept that natural humic and fulvic acids are composed of substituted monomeric species which form weak aggregates in natural waters (Peuravuori 2005; Hertkorn et al. 2006).

The more apolar, oxygen poor precursor ions required more energy to fragment (Capley et al. 2010), leading to a higher relative intensity of precursor peak remaining after fragmentation at a given energy (Fig. 3). However, contrary to our expectations, the fragmentation pattern hardly changed for any particular precursor peak over the course of the chromatographic separation (Fig. 3).

The fragmentation pattern changed little over the polarity separation. There was often a slight increase in 2CO_2 loss and sometimes an equivalent decrease in H_2O loss. Loss of CO_2 , usually the most abundant fragment, typically stayed surprisingly uniform. The increased loss of 2CO_2 was likely due to an increase in the more acidic isomers at higher retention times, as these compounds are likely to have a higher number of labile carboxylic acid groups, similar to model compound A, which had a high retention time, four such functional groups and a very large 2CO_2 loss peak (Figs. 1b, 3d).

The CID fragmentation pattern of peaks in natural organic matter is often found to be dissimilar to any purchased model

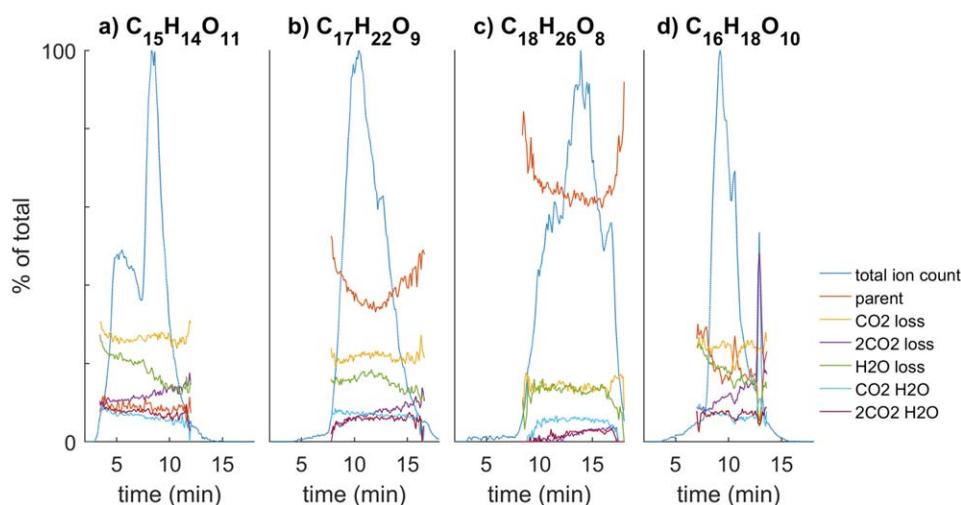


Fig. 3. Fragmentation patterns of four precursor peaks in the dataset after CID at 21% normalized energy, shown as percentage of total signal attributed to that specific precursor, over the chromatographic run. Fragment abundance data are only shown where the total ion current is above 5% of the maximum value. (a–c) $C_{15}H_{14}O_{11}$, $C_{17}H_{22}O_9$, and $C_{18}H_{26}O_8$ in the headwater stream sample, (d) $C_{16}H_{18}O_{10}$ in the SRFA sample, with the model compounds added. The model compounds upset the uniform fragmentation pattern, as seen in the other examples (a–c).

compound, as has been discussed previously (Leenheer et al. 2001; Witt et al. 2009; Capley et al. 2010; Zark et al. 2017). The most recent theory to explain this is that so many isomers are contributing to the signal that an average result is obtained—statistically described as the “central limit theorem” (Zark et al. 2017). The fragmentation pattern then takes on this average signal as the sum of all the constituent fragmentation patterns (Capley et al. 2010; Zark et al. 2017). The implication of our result is that this isomeric averaging continues even after polarity separation, so that not only is the central limit of fragmentation patterns obtained at any particular retention time, but every retention time has a similar amount of isomeric averaging and a similar result. In this case, the total number of compounds present in seawater (100,000) calculated by Zark et al. (2017) may be at least an order of magnitude too low (a result that they do not rule out).

Alternatively, our results may be explained by similar functional group chemistry of the various structural isomers leading to similar fragmentation patterns, and chemical differences that lead to the polarity distribution being present on the carbon backbone (Capley et al. 2010). This issue is difficult to resolve with our online technique. Generally, there is not enough material for MS^3 or MS^4 fragmentation at any chromatographic time to probe deeper structural differences (Leenheer et al. 2001; Capley et al. 2010). The two glycoside model compounds we analyzed (B and C) had largely different structures and different carboxylic acid abundance (0 vs. 2), but very similar retention times. Model compound A had a much longer retention time, suggesting that it is rather hydrophobic when neutralized (Supporting Information Fig. S2). The parallel retention of acidic and purely hydrophobic functionality perhaps obscures and

complicates potentially important trends in the fragmentation data. A stationary phase with a different selectivity such as hydrophilic interaction liquid chromatography may help in future work (Woods et al. 2011). However, we can state that the natural mixture is complex and yet gives a surprisingly consistent fragmentation pattern that can be significantly disrupted by simply adding in a small concentration of a known compound (Fig. 3d).

Up to this point, we have only considered the most important neutral losses (combinations of CO_2 and H_2O) because we could not unambiguously assign other atomic combinations. For example, a loss of vinyl ketene (C_4H_4O , 68.026 Da) was observed for some peaks, indicating the presence of cyclic ketone functionality (Harris et al. 1967), but this may also be due to loss of carbon suboxide (C_3O_2 , 67.990 Da) (Huber et al. 2007) from the previous precursor peak. These and other more exotic neutral losses may only be investigated with clarity with more advanced techniques that can isolate a single precursor peak (Witt et al. 2009; Brown et al. 2016). These pre-selection techniques would come with a necessary loss in time resolution, but would greatly improve the scope of the technique that we used. It is possible that individual compounds could be investigated this way based on their specific losses.

It was possible to identify some chromatographically resolved fragments from the complex mixtures without assigning them to specific precursor ions. Such peaks were a small minority compared with the typical broad peaks and were generally low in abundance. Example chromatograms are available in the Supporting Information. Particularly interesting were fragments that were only identified in seawater samples, such as $C_{10}H_{17}O_5^-$ at 17.27 min, $C_{14}H_7O_{11}^-$ at

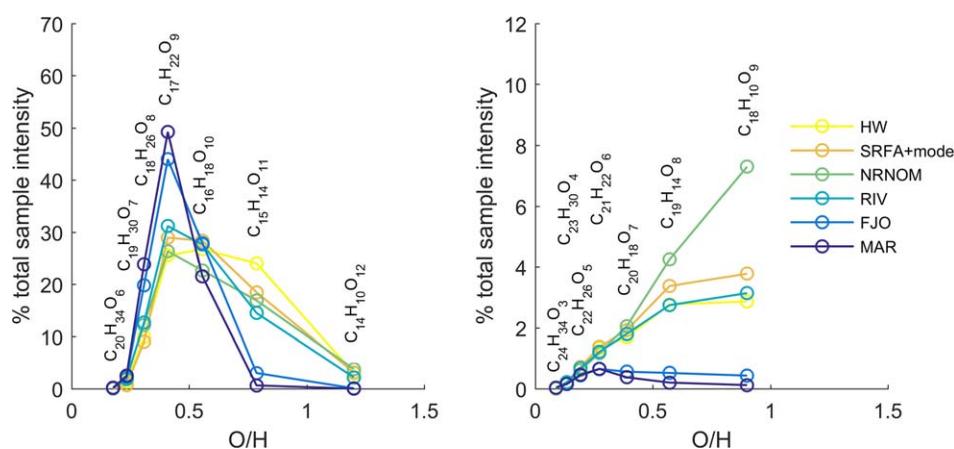


Fig. 4. Relative abundances of peaks found, integrated across the whole chromatographic run, normalized to 100% per sample. For clarity, the results are divided into two panels showing the two series.

9.46 min, and $C_{13}H_5O_5^-$ at 17.02 min, as these peaks presumably originate from molecules that are not present in terrestrial waters, and so are specific to seawater primary production or degradation products of terrestrial compounds. It was our expectation that more features like this would be visible after chromatographic separation, but in reality, the vast majority of fragment ion intensity took on the type of broad, average pattern that analytical chemists are accustomed to seeing in natural organic mixtures, with any analytical technique. We find this to be a fascinating result that indicates that the composition of natural organic matter is transformed to an ultra-complex mixture in every aquatic environment, the diversity of which is only constrained by structural possibilities and probability.

Comparison of various aquatic samples

The six samples analyzed had different abundances of the molecular masses with m/z 369 (Fig. 4). The freshwater samples had more of the highly oxidized and unsaturated molecular masses (high O/H ratio, e.g., $C_{15}H_{14}O_{11}$) compared with the samples from seawater. These highly oxidized and unsaturated molecules tend to come from terrestrial organic matter, and are known to be gradually diluted in seawater (Medeiros et al. 2016). Moreover, new supply of more saturated and reduced compounds by marine primary production may have contributed to the relative increase in these compounds in marine waters.

The polarity-dependent fragmentation pattern of these different precursor peaks was astonishingly similar in the various samples, which were taken from six completely different environments (Fig. 5), despite a roughly 100-fold variation in dissolved organic carbon (DOC) concentration and several hundreds of years of difference in aquatic residence time (Catalán et al. 2016) and therefore biogeochemical processing. It may be said that the marine samples showed greater loss of water than the terrestrial samples for more oxygen rich peaks, possibly indicating an increase in alcohol functionality, but

this trend is rather weak and not consistent for all precursors (Fig. 5). The overall fragmentation similarity suggests that the individual structural components of the isomeric mixture are either the same across aquatic environments, or that the distributions of structures are centered around similar average patterns (Reemtsma et al. 2006). It also shows that all of the samples are extremely complex to the extent that this technique cannot readily distinguish between them.

It is possible to assess the functional complexity of individual samples based on the absolute number of different fragments generated, even though most of the low abundance fragments could not be unambiguously assigned to a precursor. The samples taken from closer to the marine end-member—from the coastal end of a river, a fjord and seawater—had the highest number of fragments and for this reason, we propose these three samples had the highest functional complexity (Fig. 2). This result contrasts with recent evidence that fresher waters (i.e., headwater streams) are more complex, based on the number of assigned formulas (Mosher et al. 2015). These two results can be reconciled if either the larger molecular diversity of fresher samples is degraded into a smaller range of masses but with an increased level of structural diversity, or if subsequent new formation by indigenous primary production in marine waters causes increased functional diversity. These effects would not be measurable by broadband mass spectrometry alone.

DOM persistence in the environment and recalcitrance to bacterial degradation is related to its diversity (Kellerman et al. 2015). Deeper insight into this diversity is crucial to understand the persistence and reactivity of DOM. It has been suggested that, in oceanic waters, most of the DOM is below the concentration threshold allowing bacterial degradation due to the implied low concentration of the thousands or millions of individual compounds present (Arrieta et al. 2015). However, within the concentration range typical for freshwater environments, bacterial exploitation of DOM

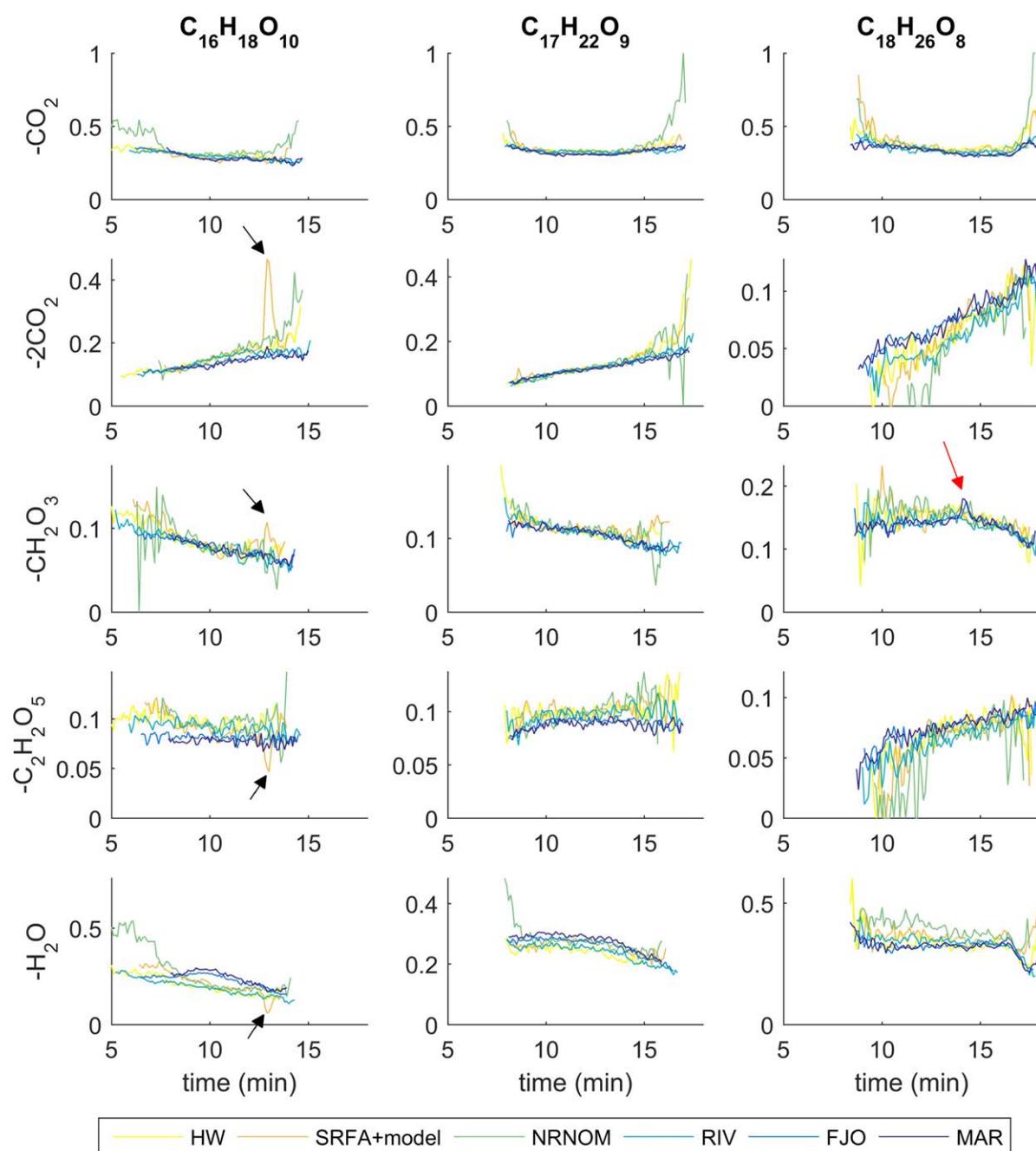


Fig. 5. Relative intensities of different fragments representing neutral losses of combinations of CO_2 and H_2O over the chromatographic run at 27% normalized CID energy from three precursor peaks (left to right) in the six samples (shown as different colors, see legend). Cooler colors represent samples with longer aquatic residence times. Data is only shown where the total ion count for the respective precursor peak exceeds 5% of the maximum value, as in Fig. 3. Features caused by model compound A are indicated by black arrows. The red arrow indicates a possible feature found in the marine samples, with specific fragment $\text{C}_{14}\text{H}_{20}\text{O}$.

does not appear to be constrained by concentration (Eiler et al. 2003), although we here demonstrate diversity of compounds similar to that in seawater. There is therefore a need for deeper information into how the huge chemodiversity of DOM translates into functional diversity such as the diversity of enzymatic pathways required for initial breakdown.

Conclusion

Natural DOM from diverse environments is too complex for HPLC-MS/MS to distinguish fragmentation patterns from individual isomers from a molecular mass. Furthermore, the fragmentation pattern obtained bears resemblance to the average signal obtained when the bulk sample is infused,

rather than chromatographically separated. The pattern is dissimilar to purchased compounds and may be an average signal of numerous structural isomers. Together, these results suggest that isomeric complexity is at least an order of magnitude greater than previously realized. Samples from diverse aquatic environments have similar levels of complexity and similar fragmentation signals, confirming that complexity is an inherent and ubiquitous feature of natural DOM. This feature of DOM should be carefully considered in all studies that use mass spectrometry to study the environmental processing of DOM without additional analysis that is sensitive to molecular structure, particularly direct infusion mass spectrometry, but also after one-dimensional chromatographic separation.

References

- Arakawa, N., and L. Aluwihare. 2015. Direct identification of diverse alicyclic terpenoids in Suwannee River Fulvic Acid. *Environ. Sci. Technol.* **49**: 4097–4105. doi:10.1021/es5055176
- Arakawa, N., L. I. Aluwihare, A. J. Simpson, R. Soong, B. M. Stephens, and D. Lane-Coplen. 2017. Carotenoids are the likely precursor of a significant fraction of marine dissolved organic matter. *Sci. Adv.* **3**: e1602976. doi:10.1126/sciadv.1602976
- Arrieta, J. M., E. Mayol, R. L. Hansman, G. J. Herndl, T. Dittmar, and C. M. Duarte. 2015. Dilution limits dissolved organic carbon utilization in the deep ocean. *Science* **348**: 331–333. doi:10.1126/science.1258955
- Brown, T. A., B. A. Jackson, B. J. Bythell, and A. C. Stenson. 2016. Benefits of multidimensional fractionation for the study and characterization of natural organic matter. *J. Chromatogr. A* **1470**: 84–96. doi:10.1016/j.chroma.2016.10.005
- Capley, E. N., J. D. Tipton, A. G. Marshall, and A. C. Stenson. 2010. Chromatographic reduction of isobaric and isomeric complexity of fulvic acids to enable multi-stage tandem mass spectral characterization. *Anal. Chem.* **82**: 8194–8202. doi:10.1021/ac1016216
- Catalán, A. N., R. Marcé, D. N. Kothawala, and L. J. Tranvik. 2016. Organic carbon decomposition rates controlled by water retention time across inland waters. *Nat. Geosci.* **9**: 501–504. doi:10.1038/ngeo2720
- Cortés-Francisco, N., and J. Caixach. 2015. Fragmentation studies for the structural characterization of marine dissolved organic matter. *Anal. Bioanal. Chem.* **407**: 2455–2462. doi:10.1007/s00216-015-8499-3
- Dittmar, T., and A. Stubbins. 2014. *Dissolved organic matter in aquatic systems*, 2nd ed. Elsevier.
- Eiler, A., S. Langenheder, S. Bertilsson, and L. J. Tranvik. 2003. Heterotrophic bacterial growth efficiency and community structure at different natural organic carbon concentrations. *Appl. Environ. Microbiol.* **69**: 3701–3709. doi:10.1128/AEM.69.7.3701-3709.2003
- Gaspar, A., M. Harir, N. Hertkorn, and P. Schmitt-Kopplin. 2010. Preparative free-flow electrophoretic offline ESI-Fourier transform ion cyclotron resonance/MS analysis of Suwannee River fulvic acid. *Electrophoresis* **31**: 2070–2079. doi:10.1002/elps.200900726
- Harris, R. L. N., F. Komitsky, and C. Djerassi. 1967. Mass spectrometry in structural and stereochemical problems. CXXXII. Electron impact induced alkyl and aryl rearrangements in α -arylidene cyclic ketones. *J. Am. Chem. Soc.* **89**: 4765–4775. doi:10.1021/ja00994a035
- Hendrickson, C. L., and others. 2015. 21 Tesla Fourier transform ion cyclotron resonance mass spectrometer: A national resource for ultrahigh resolution mass analysis. *J. Am. Soc. Mass Spectrom.* **26**: 1626–1632. doi:10.1007/s13361-015-1182-2
- Hertkorn, N., R. Benner, M. Frommberger, P. Schmitt-Kopplin, M. Witt, K. Kaiser, A. Kettrup, and J. I. Hedges. 2006. Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Acta* **70**: 2990–3010. doi:10.1016/j.gca.2006.03.021
- Huber, S. G., G. Kilian, and H. F. Scholer. 2007. Carbon suboxide, a highly reactive intermediate from the abiotic degradation of aromatic compounds in soil. *Environ. Sci. Technol.* **41**: 7802–7806. doi:10.1021/es071530z
- Kellerman, A. M., D. N. Kothawala, T. Dittmar, and L. J. Tranvik. 2015. Persistence of dissolved organic matter in lakes related to its molecular characteristics. *Nat. Geosci.* **8**: 454–459. doi:10.1038/NNGEO2440
- Koch, B. P., K.-U. Ludwigowski, G. Kattner, T. Dittmar, and M. Witt. 2008. Advanced characterization of marine dissolved organic matter by combining reversed-phase liquid chromatography and FT-ICR-MS. *Mar. Chem.* **111**: 233–241. doi:10.1016/j.marchem.2008.05.008
- Lam, B., A. Baer, M. Alaei, B. Lefebvre, A. Moser, A. Williams, and A. J. Simpson. 2007. Major structural components in freshwater dissolved organic matter. *Environ. Sci. Technol.* **41**: 8240–8247. doi:10.1021/es0713072
- Lam, B., and A. J. Simpson. 2009. Investigating aggregation in Suwannee River, USA, dissolved organic matter using diffusion-ordered nuclear magnetic resonance spectroscopy. *Environ. Toxicol. Chem.* **28**: 931–939. doi:10.1897/08-441.1
- Leenheer, J. A., C. E. Rostad, P. M. Gates, E. T. Furlong, and I. Ferrer. 2001. Molecular resolution and fragmentation of fulvic acid by electrospray ionisation/multistage tandem mass spectrometry. *Anal. Chem.* **73**: 1461–1471. doi:10.1021/ac0012593
- Longnecker, K., and E. B. Kujawinski. 2017. Mining mass spectrometry data: Using new computational tools to find novel organic compounds in complex environmental mixtures. *Org. Geochem.* **110**: 92–99. doi:10.1016/j.orggeochem.2017.05.008

- Marshall, A. G., C. L. Hendrickson, and G. S. Jackson. 1998. Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* **17**: 1–35. doi:10.1002/(SICI)1098-2787(1998)17:1 <1::AID-MAS1 >3.0.CO;2-K
- Medeiros, P. M., and others. 2016. A novel molecular approach for tracing terrigenous dissolved organic matter into the deep ocean. *Global Biogeochem. Cycles* **30**: 689–699. doi:10.1002/2015GB005320
- Mosher, J. J., L. A. Kaplan, D. C. Podgorski, A. M. McKenna, and A. G. Marshall. 2015. Longitudinal shifts in dissolved organic matter chemogeography and chemodiversity within headwater streams: A river continuum reprise. *Biogeochemistry* **124**: 371–385. doi:10.1007/s10533-015-0103-6
- Namjesnik-Dejanovic, K., and S. E. Cabaniss. 2004. Reverse-phase HPLC method for measuring polarity distributions of natural organic matter. *Environ. Sci. Technol.* **38**: 1108–1114. doi:10.1021/es0344157
- Patriarca, C., J. Bergquist, P. J. R. Sjöberg, L. Tranvik, and J. A. Hawkes. 2017. Online HPLC-ESI-HRMS method for the analysis and comparison of different dissolved organic matter samples. *Environ. Sci. Technol.* doi:10.1021/acs.est.7b04508
- Petras, D., and others. 2017. High-resolution liquid chromatography tandem mass spectrometry enables large scale molecular characterization of dissolved organic matter. *Front. Mar. Sci.* **4**: 405. doi:10.3389/fmars.2017.00405
- Peuravuori, J. 2005. NMR spectroscopy study of freshwater humic material in light of supramolecular assembly. *Environ. Sci. Technol.* **39**: 5541–5549. doi:10.1021/es050341i
- Plancque, G., B. Amekraz, V. Moulin, P. Toulhoat, and C. Moulin. 2001. Molecular structure of fulvic acids by electrospray with quadrupole time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **15**: 827–835. doi:10.1002/rcm.307
- Reemtsma, T., and A. These. 2003. On-line coupling of size exclusion chromatography with electrospray ionization-tandem mass spectrometry for the analysis of aquatic fulvic and humic acids. *Anal. Chem.* **75**: 1500–1507. doi:10.1021/ac0261294
- Reemtsma, T., A. These, A. Springer, and M. Linscheid. 2006. Fulvic acids as transition state of organic matter: Indications from high resolution mass spectrometry. *Environ. Sci. Technol.* **40**: 5839–5845. doi:10.1021/es060318c
- Riedel, T., and T. Dittmar. 2014. A method detection limit for the analysis of natural organic matter via Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **86**: 8376–82. doi:10.1021/ac501946m
- Rodgers, R. P., T. M. Schaub, and A. G. Marshall. 2005. PETROLEOMICS: MS returns to its roots. *Anal. Chem.* **77**: 20 A–27 A. doi:10.1021/ac053302y
- Saleh, F. Y., W. A. Ong, and D. Y. Chang. 1989. Structural features of aquatic fulvic acids. Analytical and preparative reversed-phase high-performance liquid chromatography separation with photodiode array detection. *Anal. Chem.* **61**: 2792–2800. doi:10.1021/ac00199a021
- Sandron, S., N. W. Davies, R. Wilson, A. R. Cardona, P. R. Haddad, P. N. Nesterenko, and B. Paull. 2017. Fractionation of dissolved organic matter on coupled reversed-phase monolithic columns and characterisation using reversed-phase liquid chromatography-high resolution mass spectrometry. *Chromatographia.* 1–11. doi:10.1007/s10337-017-3324-0
- Sleighter, R. L., Z. Liu, J. Xue, and P. G. Hatcher. 2010. Multivariate statistical approaches for the characterization of dissolved organic matter analyzed by ultrahigh resolution mass spectrometry. *Environ. Sci. Technol.* **44**: 7576–7582. doi:10.1021/es1002204
- Stenson, A. C. 2008. Reversed-phase chromatography fractionation tailored to mass spectral characterization of humic substances. *Environ. Sci. Technol.* **42**: 2060–2065. doi:10.1021/es7022412
- Tranvik, L. J., and others. 2009. Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol. Oceanogr.* **54**: 2298–2314. doi:10.4319/lo.2009.54.6_part_2.2298
- Witt, M., J. Fuchser, and B. P. Koch. 2009. Fragmentation studies of fulvic acids using collision induced dissociation Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **81**: 2688–2694. doi:10.1021/ac802624s
- Woelki, G., S. Friedrich, G. Hanschmann, and R. Salzer. 1997. HPLC fractionation and structural dynamics of humic acids. *Fresenius J. Anal. Chem.* **357**: 548–552. doi:10.1007/s002160050211
- Woods, G. C., M. J. Simpson, P. J. Koerner, A. Napoli, and A. J. Simpson. 2011. HILIC-NMR: Toward the identification of individual molecular components in dissolved organic matter. *Environ. Sci. Technol.* **45**: 3880–3886. doi:10.1021/es103425s
- Wu, Z., R. P. Rodgers, and A. G. Marshall. 2004. Two- and three-dimensional van Krevelen diagrams: A graphical analysis complementary to the Kendrick mass plot for sorting elemental compositions of complex organic mixtures based on ultrahigh-resolution broadband Fourier transform ion cyclotron resonance. *Anal. Chem.* **76**: 2511–2516. doi:10.1021/ac0355449
- Wu, C., K. Qian, C. C. Walters, and A. Mennito. 2015. Application of atmospheric pressure ionization techniques and tandem mass spectrometry for the characterization of petroleum components. *Int. J. Mass Spectrom.* **377**: 728–735. doi:10.1016/j.ijms.2014.08.019
- Zark, M., J. Christoffers, and T. Dittmar. 2017. Molecular properties of deep-sea dissolved organic matter are predictable by the central limit theorem: Evidence from tandem FT-ICR-MS. *Mar. Chem.* **191**: 9–15. doi:10.1016/j.marchem.2017.02.005

Acknowledgments

The manuscript benefitted from discussions with K. Duncan and from the comments of two anonymous reviewers. This research was made possible with grants from the Knut and Alice Wallenberg Foundation (Grant 2013.0091) and the Olsson-Borgh Foundation (Grant 130305807). Cruise JC082 from which the marine sample was collected was funded by the Natural Environment Research Council of the UK

(Grant NE/F017774/1). The headwater stream sample was provided by Stefan Löfgren (Swedish University of Agricultural Sciences) funded by the Swedish Energy Agency.

Submitted 20 August 2017

Revised 20 November 2017

Accepted 14 January 2018