



## Hydrocarbons to carboxyl-rich alicyclic molecules: A continuum model to describe biodegradation of petroleum-derived dissolved organic matter in contaminated groundwater plumes

David C. Podgorski<sup>a,\*</sup>, Phoebe Zito<sup>a,2</sup>, Anne M. Kellerman<sup>b,c</sup>, Barbara A. Bekins<sup>d,\*</sup>,  
Isabelle M. Cozzarelli<sup>e,4</sup>, Donald F. Smith<sup>c,5</sup>, Xiaoyan Cao<sup>f</sup>, Klaus Schmidt-Rohr<sup>f</sup>,  
Sasha Wagner<sup>g,h</sup>, Aron Stubbins<sup>h,6</sup>, Robert G.M. Spencer<sup>b,c</sup>

<sup>a</sup> Pontchartrain Institute for Environmental Sciences, Department of Chemistry, Chemical Analysis & Mass Spectrometry Facility, University of New Orleans, New Orleans, LA 70148, USA

<sup>b</sup> Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, FL, USA

<sup>c</sup> National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL 32310, USA

<sup>d</sup> US Geological Survey, Menlo Park, CA 94025, USA

<sup>e</sup> US Geological Survey, Reston, VA 20192, USA

<sup>f</sup> Department of Chemistry, Brandeis University, Waltham, MA 02453, USA

<sup>g</sup> Department of Earth and Environmental Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

<sup>h</sup> Department of Chemistry and Chemical Biology, Department of Marine and Environmental Sciences, Department of Civil and Environmental Engineering, Northeastern University, Boston, MA 02115, USA

### ARTICLE INFO

#### Keywords:

Hydrocarbon oxidation products  
Polar hydrocarbon metabolites  
OPAH  
Oxyhydrocarbon

### ABSTRACT

Relationships between dissolved organic matter (DOM) reactivity and chemical composition in a groundwater plume containing petroleum-derived DOM (DOM<sub>HC</sub>) were examined by quantitative and qualitative measurements to determine the source and chemical composition of the compounds that persist downgradient. Samples were collected from a transect down the core of the plume in the direction of groundwater flow. An exponential decrease in dissolved organic carbon concentration resulting from biodegradation along the transect correlated with a continuous shift in fluorescent DOM<sub>HC</sub> from shorter to longer wavelengths. Moreover, ultrahigh resolution mass spectrometry showed a shift from low molecular weight (MW) aliphatic, reduced compounds to high MW, unsaturated (alicyclic/aromatic), high oxygen compounds that are consistent with carboxyl-rich alicyclic molecules. The degree of condensed aromaticity increased downgradient, indicating that compounds with larger, conjugated aromatic core structures were less susceptible to biodegradation. Nuclear magnetic resonance spectroscopy showed a decrease in alkyl (particularly methyl) and an increase in aromatic/olefinic structural motifs. Collectively, data obtained from the combination of these complementary analytical techniques indicated that changes in the DOM<sub>HC</sub> composition of a groundwater plume are gradual, as relatively low molecular weight (MW), reduced, aliphatic compounds from the oil source were selectively degraded and high MW, alicyclic/aromatic, oxidized compounds persisted.

\* Corresponding authors.

E-mail addresses: [dcpodgor@uno.edu](mailto:dcpodgor@uno.edu) (D.C. Podgorski), [babekins@usgs.gov](mailto:babekins@usgs.gov) (B.A. Bekins).

<sup>1</sup> ORCID: DCP: 0000-0002-1070-5923.

<sup>2</sup> ORCID: PZ: 0000-0002-7011-1940.

<sup>3</sup> ORCID: BAB: 0000-0002-1411-6018.

<sup>4</sup> ORCID: IMC: 0000-0002-5123-1007.

<sup>5</sup> ORCID: DFS: 0000-0003-3331-0526.

<sup>6</sup> ORCID: AS: 0000-0002-3994-1946.

<https://doi.org/10.1016/j.jhazmat.2020.123998>

Received 27 January 2020; Received in revised form 17 August 2020; Accepted 13 September 2020

Available online 19 September 2020

0304-3894/© 2020 The Author(s).

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Organic carbon mobilized from hydrocarbons released into the environment in the form of petroleum-derived dissolved organic matter (DOM<sub>HC</sub>) poses acute risks to humans and ecosystems. While it is major oil spills such as Exxon Valdez and Deepwater Horizon that capture the headlines, there are approximately 126,000 petroleum-contaminated groundwater sites in the United States alone (NRC, 2012; Allan et al., 2012; Overton et al., 2016; Sauer et al., 1993). Plumes of mobilized organic carbon comprised of a mixture of native (background) DOM (Baker and Lamont-Black, 2001; Luzius et al., 2018), dissolved petroleum hydrocarbons (Zemo et al., 2013, 2017), and water-soluble oxygen-containing organic compounds (OCOCs) in the form of polar metabolites (microbial-derived partially oxidized hydrocarbon degradation products) (Mohler et al., 2013; Gieg and Sufita, 2002; Jobelius et al., 2011; Schrlau et al., 2017) or hydrocarbon oxidation products (Fry and Steenson, 2019) are transported downgradient where they may enter drinking water supplies or surface waters. The selective degradation and/or persistence of the DOM<sub>HC</sub> in these plumes is of interest in the context of predicting plume stability and the composition of refractory products that persist beyond the part of the plume that contains hydrocarbons (HCs).

Natural attenuation is the most accepted method for remediation of petroleum-impacted aquifers. This method utilizes in-situ processes that include biodegradation, dilution, adsorption, volatilization, and transformation at contaminated groundwater sites (EPA, 1999). Quantification of total petroleum hydrocarbons (TPH) as diesel or diesel-range organics in groundwater by solvent extraction with dichloromethane (EPA Method 3510C) followed by gas chromatography with flame ionization detection (EPA 8015 B/C) is used by regulators to monitor contaminated groundwater plumes and make decisions regarding site closure. However, neither the extraction nor detection method is specific for petroleum hydrocarbons. Rather, the extraction method is specific for all relatively non-polar analytes that are dissolved into the aqueous phase and the analytical method is specific for all analytes that have a boiling point between 170 and 430 °C, including OCOCs (Zemo et al., 2013). Interference in the accurate quantification of TPH at contaminated groundwater sites by dichloromethane-extractable OCOCs that are amenable to gas chromatography in the diesel range is well documented (Zemo and Foote, 2003). A silica gel cleanup step was proposed as a technique to remove OCOCs prior to analysis by EPA 8015 B/C (Zemo and Foote, 2003). Although the application of silica gel cleanup provides a more accurate measurement of TPH in groundwater plumes, its proposed use to make regulatory decisions at contaminated implies that hydrocarbon derived OCOCs in the form of polar biodegradation metabolites should not be considered in risk assessments.

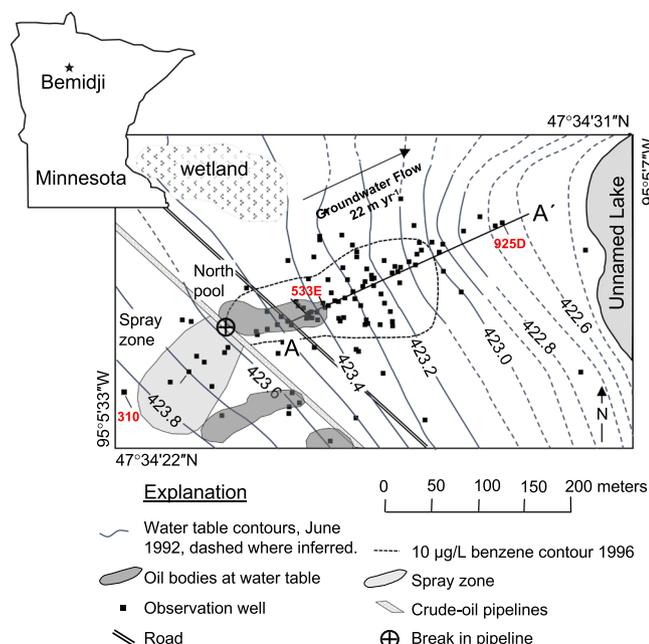
The concentration of OCOCs in groundwater at HC contaminated sites greatly exceeds that of TPH and the OCOC plume extends farther downgradient from the petroleum body than the TPH plume (Zemo and Foote, 2003; Zito et al., 2019a; Bekins et al., 2016). It is important to understand the source and chemical composition of the OCOCs in order to determine their fate and reactivity in terms of impacts on human health and the environment. There are two proposed mechanisms for the formation of the observed OCOCs in groundwater plumes. The first is that hydrocarbons are biodegraded under anoxic conditions by anaerobic pathways such as fumarate addition, oxygen independent hydroxylation, and/or carboxylation which results in the formation of partially-oxidized hydrocarbon products (Toth and Gieg, 2018; Oberding and Gieg, 2018; Fuchs et al., 2011; Heider, 2007; Callaghan, 2013). These oxygen-containing hydrocarbon products are more polar and thus more likely to partition from the oil phase into the aqueous phase. The end result is that the vast majority of the OCOCs measured as dissolved organic carbon (DOC) in the saturated zone around the petroleum body are hydrocarbon derived. Second is a microbial biosynthesis hypothesis that was proposed as a mechanism of formation for the majority OCOCs in groundwater plumes at contaminated sites (O'Reilly, 2018, 2019;

Mohler et al., 2020). This hypothesis states that OCOCs are products of microbial growth and biosynthesis, in which hydrocarbon biodegradation drives anabolic production of relatively high-MW biomolecules such as proteins, carbohydrates, lipids, and nucleic acids that are released by the microbes as OCOCs (Mohler et al., 2020). Similar to DOM that is derived from different sources in the environment, we expect oxygen-containing biosynthesis products to have analytical signatures that are distinct from hydrocarbon-derived OCOCs in the plume (Luzius et al., 2018; Kellerman et al., 2018; Drake et al., 2018).

Crude oil is a complex mixture of hydrocarbons with minor contributions from sulfur, nitrogen, and oxygen heteroatoms as well as metals such as nickel, vanadium, and iron. An understanding of the chemical composition and structure of crude oil compounds (and refined fractions) is important for determining the source, reactivity, and fate of OCOCs in groundwater plumes. In a four-part series of papers published from 1987 to 1992, Boduszynski and Altgelt showed that chemical constituents that comprise the complex mixture that is petroleum are continuous in molecular weight (MW), hydrogen deficiency, and heteroatom content (Boduszynski, 1987, 1988; Altgelt and Boduszynski, 1992; Boduszynski and Altgelt, 1992). Their "compositional continuum model" advanced the understanding of the relationships between MW, aromaticity, heteroatom content, and boiling point, which is essential for quantitative predictions of distillate fractions and behavior of heavy petroleum based on the composition of the parent crude oil. Although the applications of the compositional continuum model are used in the petroleum industry for upstream and downstream operations, this model has implications for understanding how crude oil degrades in the environment. If OCOCs are partially oxidized hydrocarbon products from microbial degradation, then the compositional continuum model of crude oil can inform us about the relationship between the molecular composition of the spilled crude oil (or refined fractions) and the composition of OCOCs that are formed from biodegradation processes.

The work presented here focuses on the water-soluble OCOCs that are operationally defined as DOM. We apply analytical techniques and concepts adapted from DOM studies to examine the source, processing, and fate of the DOM<sub>HC</sub> in the groundwater plume. One concept that will be discussed is the relationship between chemical composition and lability of DOM (Benner and Amon, 2015; Catalán et al., 2017; Mostovaya et al., 2017a; Hertkorn et al., 2008; Vähätalo et al., 2010; Guillemette et al., 2013). Similar to the compositional continuum model of crude oil proposed by Boduszynski and Altgelt, the foundation of understanding DOM reactivity is that the compounds that comprise the complex mixture that is DOM are continuous in MW, degree of saturation, and heteroatom content (Hertkorn et al., 2008). Hertkorn et al. (2008) stated that "soils, sediments, freshwaters, and marine waters contain natural organic matter, an exceedingly complex mixture of organic compounds that collectively exhibit a nearly continuous range of properties". Degradation of the bulk pool of DOM can be described by decay of compound classes and individual compounds that comprise it. Indeed, previous studies show the application of decay models to determine biodegradation kinetics of chromatographic fractions of DOM (Catalán et al., 2017) and even individual molecular formulae identified by ultrahigh-resolution mass spectrometry (UHR-MS) (Mostovaya et al., 2017a, 2017b). Accordingly, the reactivity and persistence of the DOM<sub>HC</sub> pool in groundwater plumes should arise as the result of the exponential decay of each of its individual constituents, thus of its initial molecular composition.

The National Crude Oil Spill Fate and Natural Attenuation Research Site is an ideal location to study the biodegradation of a DOM<sub>HC</sub> plume because of the extensive amount of data collected from historical and ongoing research at the site that began in 1982 (Bekins et al., 2016; Essaid et al., 2011; Baedecker et al., 1993; Fahrenfeld et al., 2014; Thorn and Aiken, 1998; Islam et al., 2016). Moreover, the site has an extensive array of groundwater monitoring wells along the core of the plume that enables sampling at high spatial resolution (Fig. 1). The documented east-northeast flow of the groundwater at an average velocity of 22 m



**Fig. 1.** Map showing the North oil pool and resulting DOM<sub>HC</sub> plume at the National Crude Oil Spill Fate and Natural Attenuation Research Site near Bemidji, MN. The samples characterized here were collected down the centerline A-A' of the plume in the direction of groundwater flow. The 310 well nest (non-plume affected), 533E (adjacent to the oil body), and 925D (toe of the plume) are labeled in red for reference.

Figure modified from Bekins et al. (2016)

y<sup>-1</sup> towards the Unnamed Lake provides the ability to utilize a space for time approach to determine rates of change in the chemical composition of DOM<sub>HC</sub> (Fig. 1) (Essaid et al., 2011). The contamination began when an oil pipeline outside the city of Bemidji, MN, USA ruptured spraying approximately 1.7 million L of light (33° API) crude oil across an area of 6500 m<sup>2</sup> (Essaid et al., 2011). The spilled oil contained 0.56% sulfur and 0.28% nitrogen with a composition of 58–61% saturated hydrocarbons, 33–35% aromatics, 4–6% resins, and 1–2% asphaltenes (w/w by saturate, aromatic, resin, and asphaltene fractionation) (Eganhouse et al., 1993). Approximately 25% of the unrecovered oil percolated through glacial outwash silt, sand and gravel to form three residual oil bodies at the water table of the underlying aquifer (Bekins et al., 2016). The vast majority of the research efforts at the site have focused on the north oil body and groundwater plume. Details about the chemical composition of the oil body (Eganhouse et al., 1993; Baedecker et al., 2011, 2018), redox conditions (Baedecker et al., 1993; Cozzarelli et al., 2016), microbial communities (Fahrenfeld et al., 2014; Beaver et al., 2016), and site hydrology (Bennett et al., 1993) are reported elsewhere (Essaid et al., 2011). The importance of biodegradation as the primary process for removal of DOC in the plume compared with dispersion, sorption and dilution has been demonstrated with modeling (Ng et al., 2015; Essaid et al., 2003), in-situ measurements (Cozzarelli et al., 2010) and by comparing volatile DOC concentrations to a conservative solute (Baedecker et al., 1993). Redox conditions are primarily methanogenic within the oil body (Bekins et al., 1999). Within the groundwater plume redox conditions are primarily iron reducing for the first 150 m (7 years) with low levels of dissolved oxygen beyond 150 m (Cozzarelli et al., 2016). Degradation rates of alkylbenzenes have been measured (Cozzarelli et al., 2010), but much less is known about the degradation of compounds in the unresolved complex mixture (Farrington and Quinn, 2015). The goal of this study is to examine the compositional controls on DOM<sub>HC</sub> reactivity across a transect of the plume to determine the persistence and fate of OCOs derived from the source.

This study utilizes complementary analytical techniques that provide

both quantitative and qualitative results including measurements of DOC concentration, excitation-emission matrix spectroscopy (EEMS), negative-ion electrospray ionization (ESI) UHR-MS, benzene polycarboxylic acid (BPCA) analyses, and nuclear magnetic resonance (NMR) spectroscopy to measure the chemical composition of a DOM<sub>HC</sub> plume. The primary objective of this work is to identify compositional pathways resulting from biodegradation of the DOM<sub>HC</sub> that can be used to determine the source, reactivity, and fate of OCOs downgradient of the oil body. Departing from the assumption that the DOM<sub>HC</sub> in the plume near the source is similar in chemical composition to the oil source, we hypothesize that all analytical measurements will show that the changes in the chemical composition of DOM<sub>HC</sub> during biodegradation are sequential and predictable with a high concentration of labile, aliphatic DOM<sub>HC</sub> adjacent to the oil body systematically decreasing to a low concentration of biorefractory DOM<sub>HC</sub> with properties similar to carboxyl-rich alicyclic molecules (CRAM). Each analytical technique is expected to show complementary molecular degradation pathways despite the selectivity or limitations of the method.

## 2. Materials and methods

### 2.1. Experimental design and sampling

Water samples were collected from 19 wells along the centerline of the north oil pool plume during the 2016–2017 field seasons. The water table is 6–8 m below the land surface and the groundwater flows east-northeast at an average velocity of 22 m y<sup>-1</sup> towards the Unnamed Lake (Fig. 1, Table 1) (Essaid et al., 2011). Wells 310B and 310E are located ~200 m upgradient from the north oil pool (Fig. 1, Table 1). The DOM from these two wells is representative of the native uncontaminated groundwater DOM at the site. Each well was purged with at least three-well volumes and samples were collected after field measurements of pH, dissolved oxygen, temperature, and specific conductance stabilized. A database of the field measurements, including replicate analyses from previous field campaigns, can be found at <https://www.sciencebase.gov/catalog/item/5910d9b2e4b0e541a03ac976?community=National+Crude+Oil+Spill+Fate+and+Natural+Attenuation+Research+Site>. All samples were filtered through 0.3 µm filters (Advantec GF-75) that were previously combusted at 450 °C > 5 h and collected in acid washed high density polyethylene bottles. The filtration step operationally-defines the water-soluble fraction (Meador and Nahrgang,

**Table 1**

The name, distance from the center of the oil body, and estimated time from the center of the oil body for each well sampled for this study.

Well	Distance from center of oil body (m)	Time from center of oil body (y)
310B	-200.1	N/A
310E	-200.5	N/A
533E	38.5	1.75
518B	55.7	2.53
518A	56.9	2.59
531A	67.5	3.07
530C	90.5	4.11
530B	91.3	4.15
530D	93.4	4.25
9315D	101.7	4.62
9315C	102.0	4.64
510	104.8	4.76
801D	123.3	5.6
801A	125.5	5.7
515B	135.7	6.17
9316C	151.1	6.87
9316A	152.7	6.94
954A	188.3	8.56
820B	210.7	9.58
820A	214.8	9.76
925D	254.4	11.56

2019) as DOM (Zito et al., 2019a, 2019b; Harriman et al., 2017; Podgorski et al., 2018; Jaggi et al., 2019; Bianchi et al., 2014; Dvorski et al., 2016). The subscript “HC” was chosen to describe DOM derived from petroleum sources because HC is commonly used for hydrocarbons, which are the major and most closely associated components of petroleum. Moreover, the DOM<sub>HC</sub> encompasses the general terms for oxygen-containing polycyclic aromatic hydrocarbons i.e., OXPAH (Koenig et al., 1983), OPAH (Nicol et al., 2001), oxy-PAH (Lundstedt et al., 2007), and the more recent term, oxyhydrocarbon (HC<sub>oxy</sub>) (Aeppli et al., 2013), that has been used to describe the residual, water insoluble fraction of weathered oil, as the source of DOM<sub>HC</sub>. Another term that is used a few times in the paper when describing plume composition and as one of the keywords is “metabolite”. This term is important because it is often used in the regulatory community to describe OCOs at petroleum contaminated groundwater sites (Zemo et al., 2013, 2017; Jobelius et al., 2011; Schrlau et al., 2017; Bekins et al., 2016; Toth and Gieg, 2018; O’Reilly Kirk et al., 2015; Gieg and Sufliita, 2005). Samples were then stored in the dark at 4 °C prior to extraction or analysis. The sampling strategy is consistent with previous studies utilizing the analytical techniques described below to measure changes in DOC concentration and DOM composition along a biogeochemical gradient (Del Vecchio and Blough, 2004; Osburn et al., 2016; Spencer et al., 2012; Stubbins et al., 2010; Koch et al., 2005; Roebuck et al., 2018; Cao et al., 2016, 2018; Tomco et al., 2019). Although replicate sample analyses were not feasible within the scope of this study due to financial restrictions, duplicate analyses for DOC were conducted (Bekins and Cozzarelli, 2017) and sampling across a gradient provided the ability to identify outliers. No outliers were identified, and the results followed expected trends.

## 2.2. Dissolved organic carbon (DOC) analyses and excitation emission matrix spectroscopic analyses

For DOC measurements, water samples were filtered through a 0.2 µm Supor® filter into a pre-combusted (550 °C > 5 h) amber glass vials and immediately acidified (pH 2, HCl) and stored in the dark at 4 °C until analysis. Previously described methods were used to measure DOC by the high temperature combustion technique with a Shimadzu TOC Vcsm analyzer (Cozzarelli et al., 2016). Separate filtered aliquots of water samples that were not previously adjusted to pH 2 were used for EEMS measurements. These filtered samples were adjusted to pH 8 to normalize for pH effects on fluorescence, particularly to avoid quenching of “humic-/fulvic-like” signatures, (Yan et al., 2013; Tfaily et al., 2011; Kulkarni et al., 2019) prior to analysis with an Aqualog® fluorometer (Horiba Scientific, Kyoto, Japan). Each sample was diluted with Milli-Q water to an absorbance value of 0.1 to reduce inner-filter effects (Ohno, 2002; Kowalczyk et al., 2003). A 10 mm quartz cuvette was used for EEMS measurements in a scan range of 240–800 nm and 0.5 s integration time. Fluorescence intensities were normalized to Raman scattering units (RSU) and dilution corrected prior to parallel factor (PARAFAC) analysis. The drEEM toolbox (tutorial and MATLAB code) was used to create and validate (residual and split-half analysis) (Harshman, 1984; Stedmon and Bro, 2008; Murphy et al., 2013) a six component PARAFAC model with 129 DOM samples. The 129 samples used for the validated model include the 19 samples reported here, samples collected from other wells at the Bemidji site, and those produced from other petroleum dissolution experiments (Podgorski et al., 2018; Zito et al., 2019b). PARAFAC is a multivariate modeling technique that statistically deconvolutes each EEM into individual fluorescent components and estimates the relative contribution of each component to the total fluorescent signal of the sample (Murphy et al., 2006, 2013). PARAFAC modeling can provide general information on source and reactivity based upon the optical character of DOM.

## 2.3. Ultrahigh resolution mass spectrometry

DOM was prepared for Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) analyses by the solid-phase extraction technique described by Dittmar et al., (2008). Briefly, filtered water samples were acidified to pH 2 and passed through a Bond Elut PPL (Agilent Technologies) cartridge. Salts were rinsed with acidified (pH 2, HCl) Milli-Q water. The stationary phase was dried with a stream of N<sub>2</sub> and the DOM was eluted with 100% MeOH (JT Baker). The final concentration of the extracts was 50 µgC mL<sup>-1</sup> and they were stored at 4 °C prior to analysis. The extraction efficiency of DOM<sub>HC</sub> across the compositional gradient of the plume originating from the north oil body at the Bemidji site by the PPL method was previously shown to be 78.4 ± 5.4% (at well 310E), 83.0 ± 4.1% (533E), 93.2 ± 12% (9315B), 83.6 ± 4.9% (925D) (Zito et al., 2019a). Direct infusion negative-ion electrospray ionization at a flow rate of 700 nL min<sup>-1</sup> with a custom built FT-ICR mass spectrometer equipped with a 21 tesla superconducting magnet was utilized for DOM analyses (Smith et al., 2018). Reproducibility of ESI UHR-MS for individual samples on multiple different instruments is reported in detail by Hawkes et al. (2020). Molecular formulae were assigned to signals >6σ RMS baseline noise with EnviroOrg software developed at the National High Magnetic Field Laboratory (NHMFL) (Corilo, 2015). A mass resolving power of 1200, 000 (m/Δm<sub>50%</sub>) was achieved at m/z 400, and the mass measurement accuracy was less than 200 ppb. Each molecular formula was classified based on stoichiometry; condensed aromatic (CA) (modified aromaticity index (AI<sub>mod</sub> ≥ 0.67), aromatic (0.67 > AI<sub>mod</sub> > 0.5), unsaturated, low oxygen (ULO) (AI<sub>mod</sub> < 0.5, H/C < 1.5, O/C < 0.5), unsaturated, high oxygen (UHO) (AI<sub>mod</sub> < 0.5, H/C < 1.5, O/C ≥ 0.5), aliphatic (H/C ≥ 1.5, N = 0) (O’Donnell et al., 2016; Šantl-Temkiv et al., 2013; Koch and Dittmar, 2006). The abundance weighted nominal oxidation state of carbon (NOSC<sub>w</sub>), H/C<sub>w</sub>, O/C<sub>w</sub> and molecular weight (MW<sub>w</sub>) were calculated based on methods described elsewhere (Riedel et al., 2012; Spencer et al., 2019; He et al., 2020).

## 2.4. Solid-state nuclear magnetic resonance spectroscopy

The major limitation with NMR for the purposes of this study is sensitivity and sample throughput. We analyzed six samples collected from the core of the plume (533E, 531A, 9315C, 801A, 954A, 925D) and one background sample (310E) by NMR spectroscopy. Solid phase extracted DOM was prepared for solid-state NMR analysis after removing the solvent (methanol) by rotary evaporation and lyophilization. Three freeze-dried samples (533E, 531A, and 9315C) were not powders (as expected for natural or background DOM); rather they were highly viscous, brown, oily substances, which made magic-angle spinning (essential in solid-state NMR experiments) challenging. These samples were therefore packed into a plastic (Kel-F) insert (Bruker Biospin), which was then packed into a 4-mm rotor. This made spinning possible and reduced the risk of a rotor crash, but introduced a small background signal near 105 ppm, which was removed after subtracting the spectrum of the neat Kel-F insert. Other samples (801A, 954A, 925D, and 310E) appeared like powders and were packed directly into a 4-mm rotor. Solid-state NMR experiments were performed on a Bruker DSX400 spectrometer at 100 MHz with a 4-mm double-resonance magic-angle spinning (MAS) probe head. Semiquantitative <sup>13</sup>C NMR spectra were collected by the multiple cross-polarization (multiCP) pulse sequence at 14 kHz MAS, (Johnson and Schmidt-Rohr, 2014) with a recycle delay of 1 s followed by eleven 1.1 ms and one 0.55 ms ramp CP periods separated by 0.5 s delays for <sup>1</sup>H relaxation and repolarization. Non-protonated carbons and mobile segments were identified by applying a period of recoupled dipolar dephasing of 68 µs to dephase magnetization of carbons with strong <sup>1</sup>H dipolar coupling, such as in immobile CH and CH<sub>2</sub> groups (Mao and Schmidt-Rohr, 2004). Signal was averaged for between 9 and 15 h for each pair of spectra. The uncertainty for NMR fractional peak areas is ±0.5% based on the signal-to-noise ratio for

signals with an area fraction <5%,  $\pm 2\%$  due to systematic errors for peaks >20%, and  $\pm 1\%$  otherwise.

### 2.5. Benzenepolycarboxylic acid analysis

The condensed aromatic carbon (ConAC) fraction of DOM was quantified and characterized using the benzenepolycarboxylic acid (BPCA) method, which oxidizes condensed aromatic structures to produce a mixture of benzenetricarboxylic acids (1,2,3-B3CA and 1,2,4-B3CA), benzenetetracarboxylic acid (1,2,4,5-B4CA), benzenepentacarboxylic acid (B5CA), and benzenhexacarboxylic acid (B6CA) molecular marker products (Dittmar, 2008). The BPCA method is a robust approach for measuring condensed aromatics in different environmental matrices, including DOM. The  $\text{DOM}_{\text{HC}}$  extracted from the samples was oxidized and the resulting BPCAs were quantified following previously described methods (Dittmar, 2008; Wagner et al., 2017a). Briefly, aliquots of DOM previously isolated by solid phase extraction ( $\sim 0.5$  mg-C equivalents) were transferred to 2 mL glass ampules and dried under a stream of argon until complete evaporation of methanol. Concentrated nitric acid (0.5 mL) was added to each ampule. Ampules were then flame-sealed and heated to 160 °C for 6 h. After oxidation, ampules were opened and dried at 60 °C under a stream of argon. The BPCA-containing residue was re-dissolved in mobile phase for subsequent analysis by high performance liquid chromatography (HPLC). BPCAs were separated and quantified on an Agilent Poroshell 120 phenyl-hexyl column using a Shimadzu HPLC system equipped with an autosampler, pump, and UV absorbance detector following procedures detailed in Wagner et al. (2017a). Individual BPCAs were quantified using calibration curves derived from a stock solution of commercially available 1,2,3-B3CA, 1,2,4-B3CA, 1,2,4,5-B4CA, B5CA, and B6CA. Coefficients of variation for replicate BPCA measurements were <5%. Sample condensed aromatic carbon concentrations were calculated using the established power relationship between ConAC ( $\mu\text{M-C}$ ) and the sum of B6CA and B5CA (nM-BPCA) using the following equation ( $n = 351$ ,  $R = 0.998$ ,  $p < 0.0001$ ;  $\text{ConAC} = 0.0891 \times (\text{B6CA} + \text{B5CA})^{0.9175}$  (Stubbins et al., 2015). BPCA ratios ( $\text{B5} + \text{B6} : \text{B3} + \text{B4}$ ), which indicate the degree of condensed aromaticity of the ConAC pool, were derived from the summed concentrations of B5CA and B6CA divided by the summed concentrations of B3CAs and B4CA.

### 2.6. Statistical analyses

Distance from center of the oil spill was converted to time using the known distance of each well from the center of the oil body and average velocity of the groundwater (Table 1). Tests of normality and exponential and linear regressions were assessed in R (RC Team, 2015). Linear models were fitted using the linear model function (lm) from the 'stats' package and exponential models were fit using the nls function which is nonlinear least squares, also from the stats package. All variables (time, DOC concentration, relative contribution of fluorescence components, relative abundance of UHR-MS formula classes, weighted averages of UHR-MS metrics, ConAC metrics, and percent NMR compounds) were normally distributed as determined by the Shapiro–Wilk test of normality ( $p > 0.05$ ). Decay of DOC concentration, fluorescence intensities, and BPCA concentrations over time was modeled using a three-parameter exponential decay function ( $C(t) = C_{\infty} + z_0 e^{-kt}$ ), which is widely used in studies assessing decay of organic matter (Lechtenfeld et al., 2014; Spencer et al., 2009). Here,  $C(t)$  = modeled value at time ( $t$ );  $C_{\infty}$  = refractory fraction at time equals infinity;  $z_0$  = biolabile fraction at  $t = 0$ ;  $k$  = rate of decay;  $t$  = time (years);  $e$  = base of natural logarithm (Table S1). The relative change over time in optical and compositional parameters was assessed using linear regressions (Table S2). The 95% confidence intervals are depicted as solid red curves on the figures and 95% prediction intervals are depicted as dashed red curves. Spearman rank correlation coefficients between molecular formula relative abundance and time were calculated in Python (Hemingway, 2017). A false discovery

rate correction was applied to correct for multiple comparisons (Benjamini and Hochberg, 1995). After the false discovery rate correction was applied, the significance level was set to  $p < 0.05$ . Molecular formulae that correlated significantly with time were then visualized in van Krevelen space (Fig. 5f).

JMP software, version 13.1.0 was used for principle components analysis (PCA). The loadings for the PCA included the normally distributed percent relative abundance scores of PARAFAC components, and metrics and compound classes identified by UHR-MS, BPCA analysis, and NMR spectroscopy. Since only a subset of samples could be analyzed by NMR spectroscopy within the scope of the project, pairwise estimation was used for the missing values in the PCA plot.

## 3. Results and discussions

### 3.1. Petroleum-derived dissolved organic carbon (DOC) degradation

The light aliphatic crude oil at the spill site has undergone weathering since the time of the spill  $\sim 40$  years ago (Baedecker et al., 2018). The relative contributions to weathering by volatilization, dissolution and biodegradation vary depending on compound properties and location in the oil body (Baedecker et al., 2018). Biodegradation and dissolution have continuously produced a  $\text{DOM}_{\text{HC}}$  plume comprised of dissolved hydrocarbons, partially-oxidized hydrocarbon metabolites, and hydrocarbon oxidation products that extends approximately 340 m downgradient from the source to a lake (Thorn and Aiken, 1998; Islam et al., 2016; Eganhouse et al., 1993). Fig. 2 and Table S3 show DOC concentrations measured for the 19 sampled wells as a function of time from the center of the oil body. A significant, exponential decrease in DOC concentration is observed along the  $\text{DOM}_{\text{HC}}$  plume transect (Fig. 2;  $p < 0.001$ ) (Essaid et al., 2003; Cozzarelli et al., 2010). The concentration of DOC in the wells directly adjacent to the oil body is  $31.1 \text{ mg L}^{-1}$ , decreasing to  $2.26 \text{ mg L}^{-1}$  in well 925D located 254 m (11.6 y) downgradient from the center of the source. The three parameters for DOC concentration vs. time are  $C_{\infty} = 0.42$ ,  $z_0 = 59.91$ , and  $k = -0.40$ . Most of the decrease occurs in the interval 0–7 years, which lies in the iron-reducing portion of the plume. The estimated degradation rate of  $-0.40 \text{ y}^{-1}$  estimated here is similar to previously modeled DOC decay estimates at the Bemidji site of  $0.46 \text{ y}^{-1}$  (Ng et al., 2015). Degradation rates of individual gas chromatography amenable compounds are reported in detail elsewhere (Cozzarelli et al., 2010; Eganhouse et al., 1996). Collectively, these results indicate that the complex mixture of  $\text{DOM}_{\text{HC}}$  is comprised of labile, semi-labile and persistent fractions

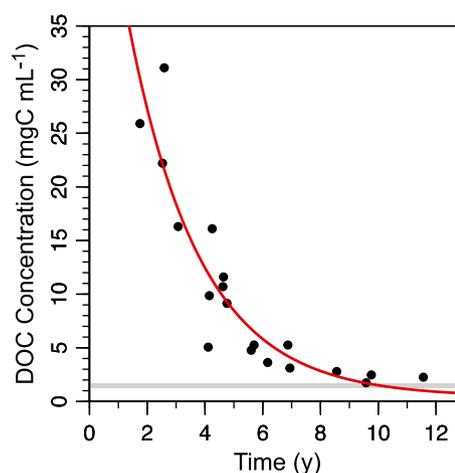


Fig. 2. DOC concentration in the groundwater plume vs. time from the center of the oil body fit to a three-parameter exponential decay model (red line). The gray shaded area represents the range of DOC concentrations measure in background wells 310B and 310E.

(Sleighter et al., 2014) providing evidence that there is a gap in knowledge about the compositional controls associated with DOC removal at hydrocarbon contaminated sites. Therefore, we will use samples from this site to study the compositional changes of the DOM<sub>HC</sub> linked to biodegradation in a petroleum-contaminated aquifer.

### 3.2. Petroleum-derived fluorescent dissolved organic matter

We previously reported optical measurements used to examine the transformation of DOM<sub>HC</sub> originating from the north oil pool at the Bemidji research site (Podgorski et al., 2018). Here, we limit the discussion to how the PARAFAC components relate to the other analytical measurements. Fig. 3 shows the six components (C1-C6) identified from the PARAFAC model and their relative contributions as a function of time from the center of the oil body. This PARAFAC model was previously published in (Podgorski et al. (2018)). There, each component is described in detail including OpenFluor matches with components derived from non-petroleum DOM sources (Podgorski et al., 2018). The model can be accessed in OpenFluor (Murphy et al., 2014). For the sake of highlighting DOM<sub>HC</sub> degradation pathways for the compounds in the unresolved complex mixture that fluoresces, the components are aligned left to right from relatively shortest to longest wavelength (i.e., lowest to highest ex./em. maxima) (C1 (<250, 280/306 nm) < C5 (275/325 nm) < C2 (<250, 285/375 nm) < C4 (305/416 nm) < C3 (<250, 305/437 nm) < C6 (265, 365/474 nm)). The bottom panels in Fig. 3 show the changes in the percent relative contribution of each PARAFAC component as a function of estimated travel time from the center of the oil source. The observed changes in the FDOM fraction of the DOM<sub>HC</sub> plume provide the first qualitative indication of the compositional and structural continuum. Each component exhibits a continuous increase or decrease in relative contribution ( $p < 0.001$  or  $0.01$ ; Fig. 3, Table S3). Moreover, the percent relative contribution of all components at the toe of the plume, with the exception of C5, do not converge with values measured in the background 310 wells (Fig. 3, Table S3). Although C1 does not completely reach background values measured at the 310 wells (0.00%), the value measured in well 925D at the toe of the plume approaches zero (0.13%). These results provide a preliminary indication that C1 and C5 are comprised of labile compounds, while the compounds in C2, C3, C4, and C6 are relatively persistent.

Fig. 4a-f and Table S3 show the fluorescence intensity of each component in Raman Scattering Units (RSU) as a function of time from

the oil body. Unlike the analysis of relative contributions reported in Fig. 3, examining the fluorescence intensity of each component provides a semi-quantitative perspective. Each component exhibits an exponential decrease in fluorescence intensity as a function of time from the oil body. The rate of decay for each component is proportional to its ex./em. maxima with the shortest wavelengths exhibiting the fastest rates of decay (Table S1). C1 and C5, the components with the shortest ex./em. maxima have the fastest rate of decay at  $0.44 \text{ y}^{-1}$  and  $0.43 \text{ y}^{-1}$ , respectively. The decay rates for C1 and C5 are both faster than that for the total DOC pool (Table S1). The rate of decay for C2 is  $0.40 \text{ y}^{-1}$ , equal to the modeled value for decay of DOC. Both C4 and C3 have the same rate of decay  $0.36 \text{ y}^{-1}$ . C6, the component with the longest ex./em. maximum, has the slowest rate of decay at  $0.30 \text{ y}^{-1}$ .

The composition of even the most complex mixture is still simply a sum of its parts. Fig. 4f-l and Table S1 show fluorescence intensity divided by DOC concentration as a function of time from the oil body. These plots compare the rate of decay for each component to the rate of DOC degradation to understand which components are comprised of labile, semi-labile, and persistent DOM<sub>HC</sub> compounds. The relationship between the decay rates of fluorescence intensity and DOC concentration are best described by linear models with the slope as an indicator of lability or persistence. The slopes obtained from dividing C1 and C5 by DOC concentration ( $p < 0.001$ ,  $r^2 = 0.86$  and  $p < 0.001$ ,  $r^2 = 0.87$  respectively) indicate that the fluorescence intensity of these components decreases much faster than DOC concentration (Fig. 4g-h). This result indicates that C1 and C5 are comprised of the most labile compounds in the DOM<sub>HC</sub> mixture, in agreement with previous reports on the labile nature of DOM compounds with similar spectral features (Fellman et al., 2010; Yamashita and Tanoue, 2003; Davis and Benner, 2007). Relative to C1 and C5, the slopes for two of the components with longer wavelengths, C2 and C3, are shallower ( $p < 0.001$ ,  $r^2 = 0.72$  and  $p < 0.05$ ,  $r^2 = 0.21$  respectively) (Fig. 4i and k). Although the fluorescence intensity of C2 and C3 decay faster than DOC concentration, they are still relatively slower than C1 and C5. Therefore, we can consider the DOM<sub>HC</sub> compounds that comprise C2 and C3 as semi-labile in this plume. C4 is another component that could be considered to be comprised of DOM<sub>HC</sub> compounds that are semi-labile (Fig. 4j). The rate of decrease in the fluorescence of compounds associated with C4 is approximately the same as DOC concentration. Conversely, the compounds that comprise C6 degrade at a slower rate than DOC concentration ( $p < 0.01$ ,  $r^2 = 0.36$ ) (Fig. 4l). This result indicates that these

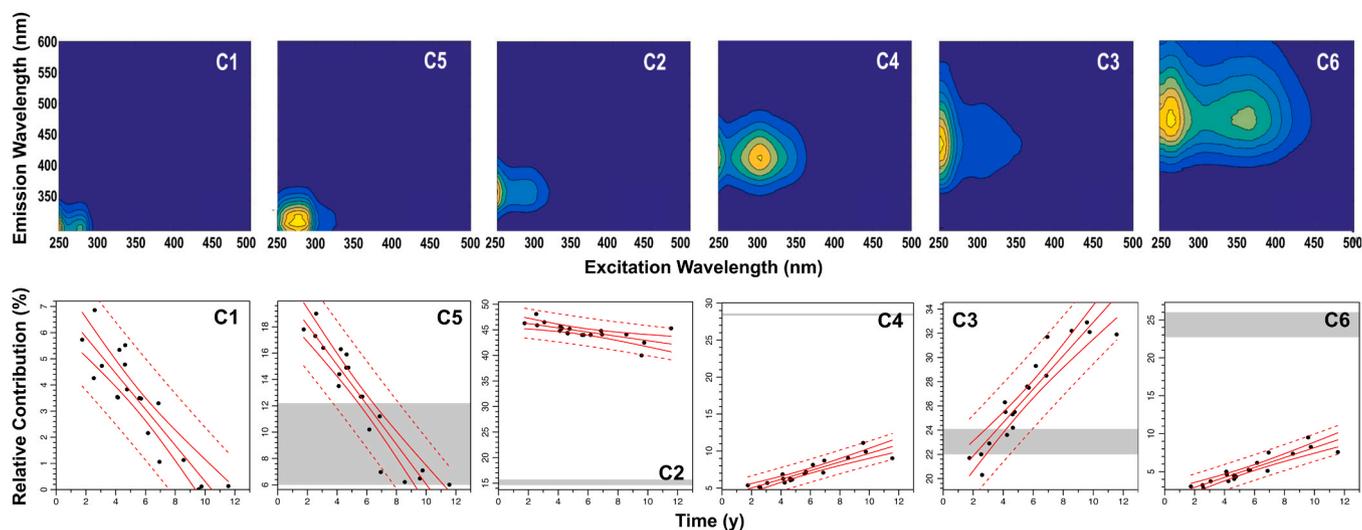
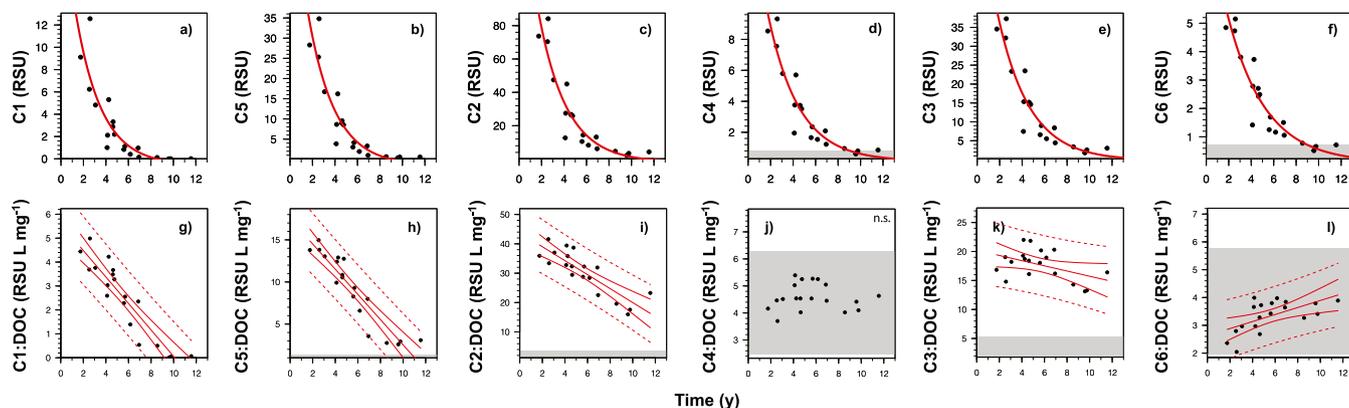


Fig. 3. Six fluorescent DOM components identified by PARAFAC analysis in order of shortest to longest wavelength (top panels). Relative contribution of each component as a function of time from the oil body with shaded gray regions representing the range of values measured for DOM collected from background wells 310B and 310E (bottom panels). The solid line in the middle shows the linear regression, 95% confidence intervals are depicted as solid red curves, and 95% prediction intervals are dashed red curves.



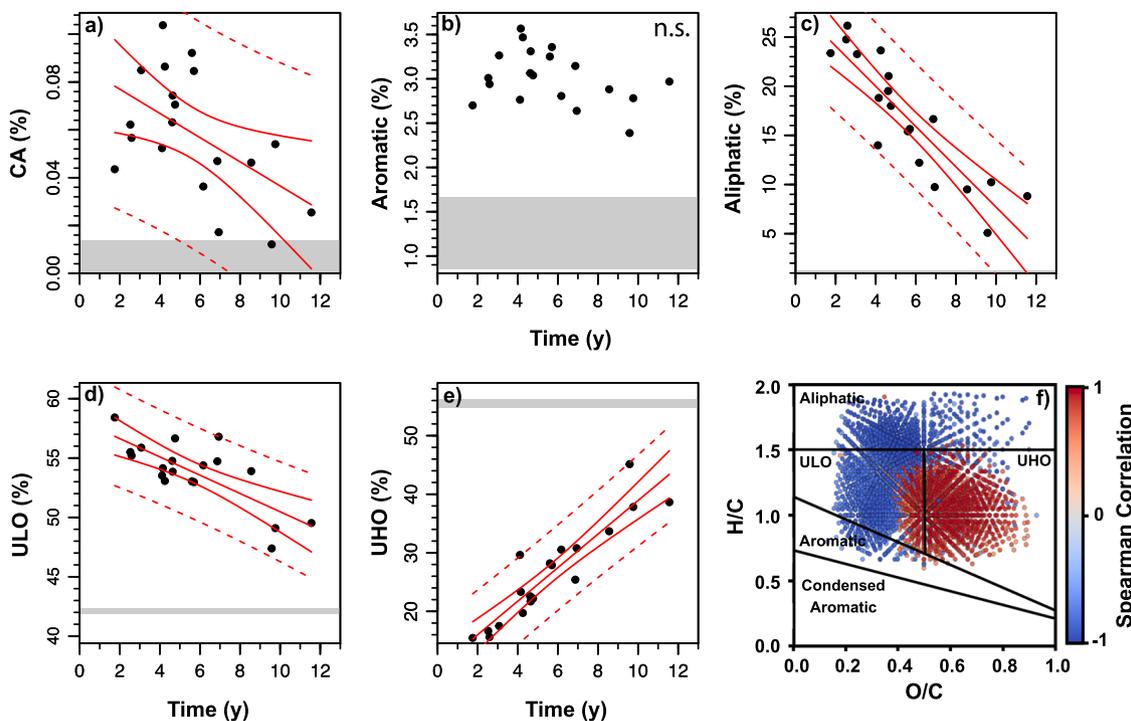
**Fig. 4.** Fluorescence intensity in Raman scattering units of components C1-C6 as a function of time from the oil body fit to the three-parameter exponential decay model (a–f). Fluorescence intensity of each component divided by DOC concentration as a function of time from the oil body (g–l). The solid line in the middle shows the linear regression, 95% confidence intervals are depicted as solid red curves, and 95% prediction intervals are dashed red curves. The shaded gray regions show the range of values measured for DOM collected from background wells 310B and 310E.

DOM<sub>HC</sub> compounds will persist the farthest downgradient. As such, the underlying compounds that comprise each component (i.e., the parts of the total fluorescent DOM<sub>HC</sub> pool) can be ranked on a scale from most labile to persistent as follows; C1, C5, C2, C3, C4, C6. The biolability of the fluorescent compounds that comprise the DOM<sub>HC</sub> in this plume is almost exactly in line with a continuum based on the ex./em. maxima of the components.

### 3.3. Molecular composition of petroleum-derived dissolved organic matter

Here we examine the fraction of DOM<sub>HC</sub> that is accessible by solid-phase extraction using PPL and subsequent negative-ion electrospray (ESI) mass spectrometry. The chemical information obtained by (–) ESI is an important part of this study because the technique is selective for

OCOCs. Fig. 5 (bottom-right) shows a visual representation of molecular formula classifications in van Krevelen (vK) space. For the purposes of this study, we highlight the degradation pathways with compound classification regions rather than focusing on nomenclature (Kim et al., 2003). In the broadest sense we examine changes in degree of saturation where condensed aromatic (CA) < aromatic < unsaturated low oxygen (ULO) < aliphatic, and oxygen content, ULO < unsaturated high oxygen (UHO). Since DOM is a complex mixture of compounds, the subdivision of vK space can yield many different regions or classifications. Thus, changes in the composition by biogeochemical processes can be selective for compounds with similar chemical composition, giving the perception of a uniform shift in vK compositional space. The defined regions provide a mechanism to track compositional degradation pathways in vK space as changes either in relative abundance or percentage



**Fig. 5.** Change in the percent relative abundance of the molecular formula classes identified by (–) ESI UHR-MS as a function of time from the center of the oil body (Table S1). Shaded gray regions indicate the range of values measured for DOM collected from background well 310B and 310E. The solid line in the middle shows the linear regression, 95% confidence intervals are depicted as solid red curves, and 95% prediction intervals are dashed red curves. Spearman correlation plot shows the correlation between time and molecular formulae in the DOM<sub>HC</sub> plume obtained by UHR-MS (bottom-right).

of assigned formulae that “shift” in and out of these regions.

Fig. 5 shows the change of the percent relative abundance of the assigned formulae in the groundwater plume as a function of estimated travel time from the oil body. The abundance of formulae classified as CA (0.01–0.10%) or aromatic (2.39–3.57%) between the source and the toe of the plume is very low because of the composition of the light, aliphatic-rich crude oil that serves as an oil source for the DOM<sub>HC</sub> plume (Fig. 5a and b). Although the crude oil that was spilled at the Bemidji site is comprised of 33–35% aromatics, the aromatics in a light crude oil are relatively small (1–3 ring) and highly alkylated (Thorn and Aiken, 1998; Eganhouse et al., 1993). Moreover, these small, highly alkylated aromatics determined by separation of the crude oil into saturates, aromatics, resins, and asphaltenes (SARA) may not fall within the compositional constraints to be classified as CA or aromatics in DOM (Koch and Dittmar, 2006). Rather, partially oxidized analogues of the SARA defined aromatics from a light, aliphatic crude oil are mostly classified as ULO or even aliphatics due to their high degree of saturation even though they may contain 1–3 ring aromatic cores (Zito et al., 2019b). Thus, contributions from formulae classified as CA and aromatic are minimal, yet there is a significant decrease in the abundance of CAs across the plume gradient (Fig. 5a;  $p < 0.05$ ;  $r^2 = 0.28$ ) whereas the decrease in aromatics with time was not significant (Fig. 5b; Table S1). The degradation of the small aromatic compounds over time is consistent with previous reports in the literature for both natural organic matter and petroleum in the environment (Dvorski et al., 2016; Anderson et al., 1998; Frontera-Suau et al., 2002; Hazen et al., 2016).

Compared with the CA and aromatic classes, it is easier to discern the

changes in the % relative abundance of the aliphatics, ULO and UHO classes as the trends are smooth and continuous along the biogeochemical gradient (Fig. 5c–e;  $r^2 = 0.78, 0.58, \text{ and } 0.85$ , respectively). Fig. 5c–d and Table S4 show that the maximum values of aliphatics (26.2%) and ULO (58.4%) are in the wells adjacent to the oil body. We expect high values for saturated and reduced classes of DOM compounds because of the aliphatic-rich composition of the hydrocarbon pool that is the source of carbon for the DOM<sub>HC</sub> plume. There is a significant decrease in the relative abundance of these two classes downgradient of the oil source to minimum values of 5.08% (aliphatics) and 47.4% (ULO) (Fig. 5c and d);  $p < 0.001$ ). This change along the gradient corresponds to a decrease of 80.6% and 18.8% for each class, respectively. Conversely, the UHO class has a minimum relative abundance of 15.5% adjacent to the oil source and exhibits a significant increase to 45.1% downgradient (Fig. 5e;  $p < 0.001$ ). These values correspond to a 191% increase in the abundance of the UHO class along the plume. The plots of each class shown in Fig. 5 indicate that with the exception of the CA class, the relative abundances of compounds that comprise each class at the toe of the plume (925D) do not converge with those measured in the 310 background wells.

Fig. 5f provides another visualization of the changes in chemical composition across the biogeochemical gradient using a Spearman correlation with estimated travel time as the independent variable. The Spearman correlation shows a negative correlation between relatively low O/C and high H/C compounds with travel time from the oil source. The composition of the plume shifts downgradient from the source to relatively high O/C, low H/C compounds. The chemical reaction

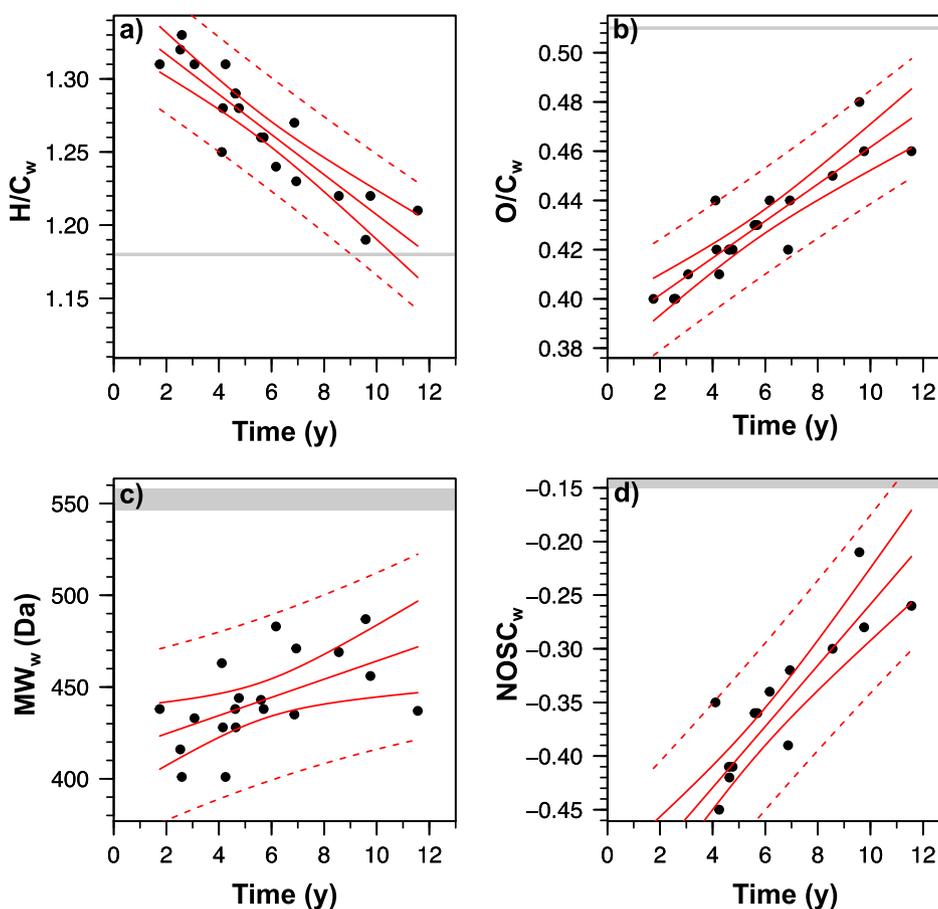


Fig. 6. The relationship between the weighted average of (a) H/C, (b) O/C, (c) MW, and (d) NOSC of the DOM<sub>HC</sub> in the plume as a function of time from the center of the oil body. Shaded gray regions indicate the range of values measured for DOM collected from background well 310B and 310E. The 95% confidence intervals are depicted as solid red curves and 95% prediction intervals are dashed red curves.

pathways responsible for these shifts are reported by Kim et al. (2003) as demethylation, dehydrogenation, and oxidation (Kim et al., 2003). The corresponding exponential decrease in DOC concentration along the gradient indicates the existence of two fractions, one behaving as labile (Fig. 5f, blue) and the other as persistent (Fig. 5f, red).

Additional values such as relative abundance weighted average  $H/C_w$ ,  $O/C_w$ , molecular weight ( $MW_w$ ), and nominal oxidation state of carbon ( $NOSC_w$ ) provide insight into the biodegradation of the  $DOM_{HC}$  in the plume. Fig. 6a–b and Table S4 show the weighted average values  $H/C_w$  and  $O/C_w$  across the plume transect. The compositional trend across the plume transect shows a decrease in saturation of the compounds that comprise the pool of DOM (Fig. 6a). The maximum  $H/C_w$  is 1.33 for the DOM adjacent to the relatively aliphatic oil source. The minimum  $H/C_w$  value of 1.19 is approximately 200 m downgradient, which corresponds to a significant decrease of 10.5% in  $H/C_w$  along the flow path (Fig. 6a;  $p < 0.001$ ,  $r^2 = 0.82$ ). Conversely, the minimum  $O/C_w$  values (0.40) are in the wells adjacent to the reduced oil source (Fig. 6b). A maximum value of 0.48 is measured downgradient, corresponding to a significant increase of 20% in  $O/C_w$  across the plume compositional gradient (Fig. 6b;  $p < 0.001$ ,  $r^2 = 0.81$ ). Like the  $H/C_w$  values, the increase in the  $O/C_w$  values along the gradient is apparently linear.

Fig. 6c and d shows plots of the  $MW_w$  and  $NOSC_w$  as a function of time from the oil source. The minimum  $MW_w$  value ( $m/z$  401 Da) is near the source while the maximum value of 487 Da occurs downgradient (Fig. 6c, Table S4). This change across the gradient corresponds to a significant increase of 21.4% in the  $MW_w$  of the compounds that comprise the pool of  $DOM_{HC}$  (Fig. 6c;  $p < 0.05$ ,  $r^2 = 0.30$ ). The increase in the  $MW_w$  along the gradient with a corresponding decrease in DOC concentration indicates that the low MW compounds are selectively degraded and relatively high MW compounds such as those derived from resins and asphaltenes persist. The persistent products may also be relatively higher in molecular weight than parent compounds from the source due to the incorporation of additional oxygen heteroatoms as a result of aqueous phase degradation. The minimum  $NOSC_w$  value of  $-0.49$  reflects the compounds near the source (Fig. 6d) and  $NOSC_w$  increases to  $-0.21$  downgradient from the oil source, corresponding to a significant increase of 46.9% along the gradient (Fig. 6d, Table S4;  $p < 0.001$ ,  $r^2 = 0.83$ ). These values indicate a pool of reduced DOM is produced from the oil source and this pool of carbon steadily becomes more oxidized downgradient. Collectively, the information obtained from examining the chemical composition by (–) ESI UHR-MS indicates that the aliphatic-rich, reduced oil source in the aquifer produces DOM that is skewed to low MW, reduced, aliphatic compounds. The microbes responsible for decreasing the DOC concentration along the gradient are selectively degrading the low MW, reduced, aliphatic compounds and leaving heavier MW, relatively aromatic/alicyclic, highly oxygenated compounds to persist downgradient. This result is consistent with previous reports for non-petroleum derived DOM (Kellerman et al., 2015; Spencer et al., 2014; Heslop et al., 2019; D'Andrilli et al., 2015) and  $DOM_{HC}$  (Harriman et al., 2017).

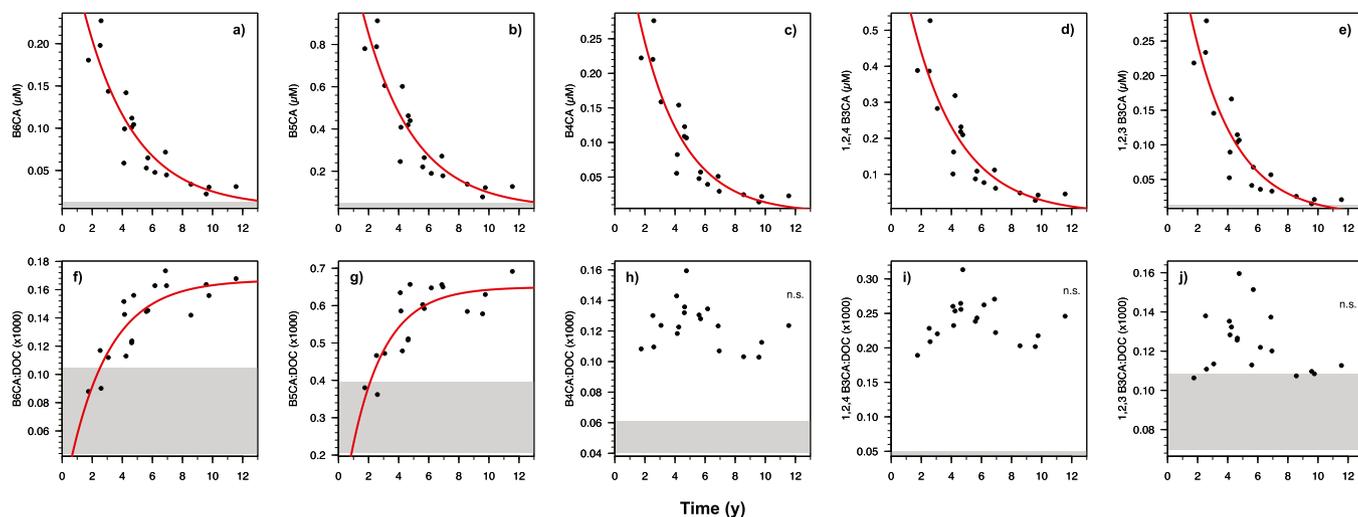
### 3.4. Distribution of condensed aromatics in petroleum-derived dissolved organic matter

We determined the transect chemistry of the condensed aromatic portion of  $DOM_{HC}$ , here termed ConAC, using the BPCA molecular markers. BPCAs are produced from condensed aromatic structures deriving from both petrogenic and pyrogenic organic matter sources (Wagner et al., 2017a; Hindersmann and Achten, 2017; Chang et al., 2018). Although the current study focuses upon the decadal evolution of DOM within a subterranean oil plume, we must also note that initial remediation of the spill ~40 years ago included the burning off of residual surface oil. Details surrounding the burn off event are scarce, therefore resulting pyrogenic inputs to the subsurface environment are unknown. As such, here we use BPCA data to track compositional

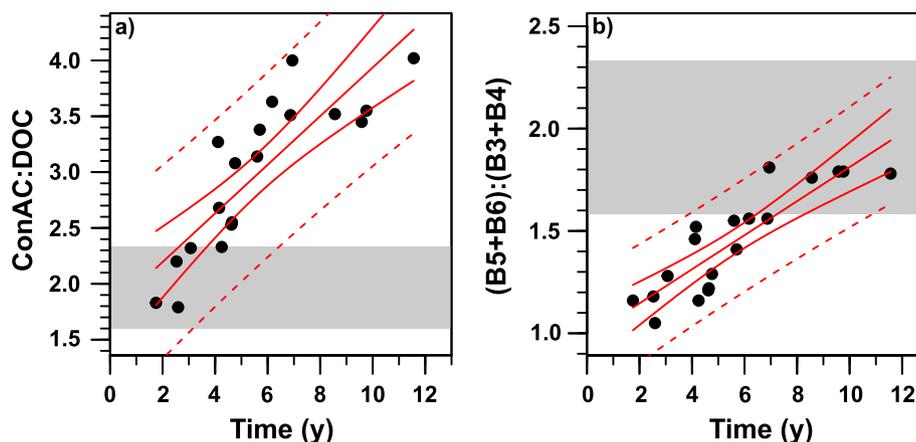
changes in ConAC along the plume transect, regardless of its origin. Overall, we did not expect the plume to contain significant quantities of ConAC, given the largely aliphatic composition of the crude oil source. Fig. 7a–e and Table S5 show values for the total concentration of ConAC and individual BPCAs measured in each well. The highest ConAC concentration ( $46.3 \mu\text{M-C}$ ) was observed near the oil source and the lowest concentration ( $5.0 \mu\text{M-C}$ ) was observed downgradient, corresponding to an 89% decrease in [ConAC] across the gradient (Table S5). The concentration of individual BPCAs decreased significantly along the oil plume, where losses  $\geq 70\%$  were observed (Fig. 7a–e, Table S5). The rates of degradation for each individual BPCA almost directly align to the number carboxylic acid groups and are always less than that of the bulk DOC (Fig. 7a–e, Table S5). The degradation rates of 1,2,3-B3CA and 1,2,4-B3CA are  $-0.34$  and  $-0.31 \text{ y}^{-1}$ , respectively (Table S1). The rate of degradation of B4CA is the same as 1,2,3-B3CA at  $0.34 \text{ y}^{-1}$  (Table S1). B5CA and B6CA exhibit the slowest rates of degradation at  $0.30 \text{ y}^{-1}$  (Table S1). These results are consistent with previous reports of biodegradation of both low- and high MW polycyclic aromatic hydrocarbons (Kanaly and Harayama, 2000; Simarro et al., 2013). Moreover, it was shown that low molecular weight organic acids enhance biodegradation of high MW PAHs (Sivaram et al., 2019). The  $DOM_{HC}$  plume at the Bemidji spill site contains large quantities of organic acids that would enhance degradation of relatively large PAH analogues that partition into the aqueous phase (Thorn and Aiken, 1998; Cozzarelli et al., 1994). These biodegradation processes are likely responsible for the decrease in ConAC concentration downgradient. However, we expect the higher molecular-weight fractions of ConAC to be more persistent downgradient.

Although concentrations of ConAC and DOC decreased downgradient, the proportion of condensed aromatic carbon in bulk DOC (ConAC:DOC) increased significantly along the transect (from 1.79% to 4.02%; Fig. 8a;  $p < 0.001$ ,  $r^2 = 0.72$ ). The composition of ConAC also varied along the gradient, becoming enriched in more highly condensed aromatic structures downgradient, as indicated by higher (B5+B6):(B3+B4) ratios in wells further from the oil source (Fig. 8b). Taken together, these results indicate that larger, condensed aromatic compounds are more biorefractory than smaller, less condensed aromatics and aliphatic compounds which comprise most of the freshly produced DOM from the oil source. As the DOC concentration decreases downgradient, aliphatic DOM is degraded at a faster rate than condensed aromatic DOM. The relative bio-recalcitrance of BPCAs derived from ConAC is consistent with previous observations for this condensed aromatic OM fraction in soils (Kuzyakov et al., 2014). Although ConAC is a small fraction of total DOC (2–4%) across the oil plume, these data corroborate results from fluorescence and ultrahigh-resolution mass spectrometry analyses. Specifically, that relatively small compounds are biodegraded while large conjugated aromatics persist downgradient.

The persistence of the highly condensed aromatic structures downgradient of the oil body was further examined by comparing the degradation rate of each BPCA with the rate of DOC degradation (Fig. 7f–j). There is no significant trend observed in the 1,2,3-B3CA, 1,2,4-B3CA, and B4CA as a function of time from the oil body (Fig. 7h–j). This result indicates that the degradation rate of those three BPCA markers are similar to the rate of DOC degradation. Interestingly, the trends observed in 1,2,3-B3CA:DOC, 1,2,4-B3CA:DOC, and B4CA:DOC are similar to the % relative abundance of the aromatics class measured by FT-ICR MS (Fig. 5b). It was previously reported that measurements by (–) ESI FT-ICR MS underrepresent ConAC, likely due to bias in ionization efficiency against this fraction of DOM (Wagner et al., 2017b). However, the similarities in trends observed between the BPCAs and relative abundance of aromatic compounds measured by FT-ICR MS support the idea that there is some overlap in analytical windows between the two techniques (Wagner et al., 2017b). These results also show that crude oil and/or fractions may be good candidates to study future hypotheses about the overlap between BPCA and FT-ICR MS measurements.



**Fig. 7.** Concentration of each BPCA in  $\mu\text{M-C}$  as a function of time from the oil body fit to the three-parameter exponential decay model (a–e). Ratio of each BPCA divided by DOC concentration ( $\mu\text{M-C}$ ) as a function of time from the oil body fit to the three-parameter exponential model (f–j). The shaded gray regions show the range of values measured for DOM collected from background wells 310B and 310E.

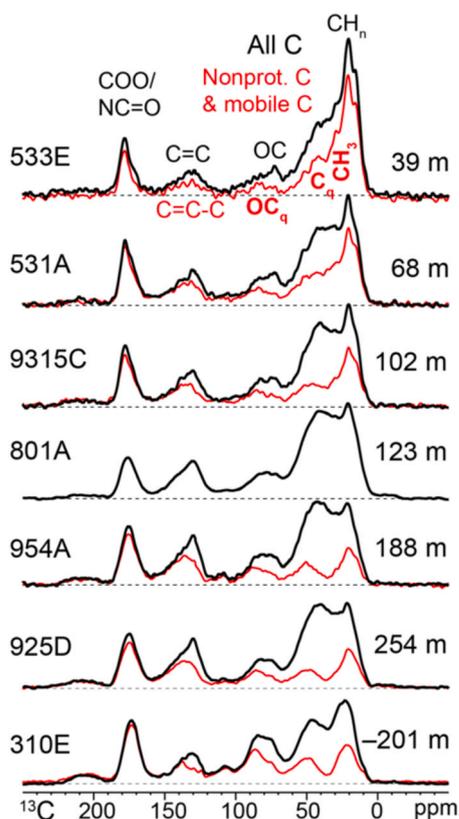


**Fig. 8.** Changes in the (a) percent ConAC:DOC and (b)  $(B5+B6):(B3+B4)$  as a function of time from the center of the oil body. Shaded gray regions indicate the range of values measured for DOM collected from background well 310B and 310E. The 95% confidence intervals are depicted as solid red curves and 95% prediction intervals are dashed red curves.

Unlike the 1,2,3-BCA, 1,2,4-BPCA, and B4CAs, the concentration of B5CA and B6CAs divided by DOC concentration result in a significant trend that can be modeled with the three-parameter exponential function (Fig. 7f and g). The results show that B5CA and B6CA degrade more slowly than the DOC pool and the curve flattens at approximately 6–7 years downgradient from the oil body. Moreover, the modeled data indicate that the B5CA and B6CA from the plume will never converge with the groundwater with  $C_{\infty} = 0.17$  and  $0.65$ , respectively. The significance of this result is that it shows that B5CA and B6CAs are conservative tracers of the contaminated plume. This is the first time that conservative tracers have been identified for petroleum-contaminated plumes. Although the source is a light aliphatic crude oil, it still contains relatively large (“heavy”) compounds (1–2% asphaltenes w/w). Partially oxidized products from the heavy end of the compositional continuum of this crude oil are the most likely source of precursors for B5CA and B6CA. Future studies will examine the utility of the BPCA method to trace  $\text{DOM}_{\text{HC}}$  produced from petroleum sources with different initial compositions in freshwater and marine ecosystems. We posit that heavy oils and fractions will produce the most biorefractory  $\text{DOM}_{\text{HC}}$  and therefore be enriched in the large BPCAs while  $\text{DOM}_{\text{HC}}$  produced from gasoline (for example) will be depleted in large BPCAs.

### 3.5. Structure of petroleum-derived dissolved organic matter

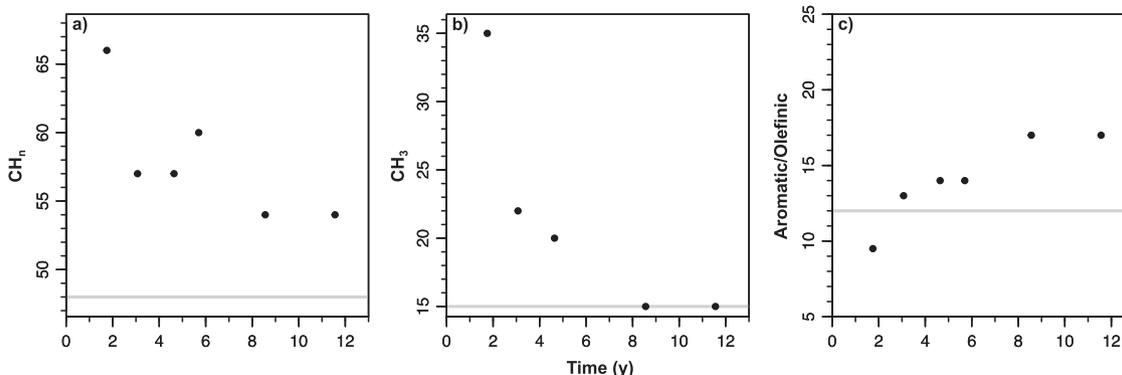
The steep biogeochemical gradient created from biodegradation of DOM produced predominantly from a single source enables us to study the changes in structure of  $\text{DOM}_{\text{HC}}$  by semi-quantitative  $^{13}\text{C}$  NMR spectroscopy. Although the NMR data set lacks high spatial resolution (i. e., sampling of plume), we can project the continuum observed by DOC, EEMs, UHR-MS, and BPCA measurements to understand the changes in the molecular structure of  $\text{DOM}_{\text{HC}}$  across the compositional gradient. Fig. 9 shows the solid-state  $^{13}\text{C}$  NMR spectra, and corresponding spectra after dipolar dephasing, of the seven DOM samples. Notably, the spectral features of DOM downgradient from the source become progressively more like those of the background DOM (310E). In addition, prominent signals from carboxyl, methyl, quaternary C ( $\text{C}_q$ ) as well as quaternary C bonded to O ( $\text{OC}_q$ ) are consistent with the presence of CRAM as previously documented in the Rifle aquifer and other DOM samples (Cao et al., 2018; Cao, 2014). Table S5 shows the percentages of ketone, COO/amide, aromatic/olefinic, (hemi)acetal, (hemi)ketal, alcohol/ether, and alkyl (or aliphatic (Thorn and Aiken, 1998) carbon and subfraction of methyl groups) measured in each sample. Because the rotor crashed during the measurement of the spectrum of the 801A sample after dipolar dephasing the NMR measurements for 801A are not



**Fig. 9.** Solid-state  $^{13}\text{C}$  NMR spectra (thick black lines) of 7 DOM samples taken at various distances from the oil source as indicated on the right. The bottom spectrum is the background reference. Thinner red lines: corresponding spectra of C not bonded to H and from mobile segments (except for sample 801A, because the sample rotor broke during measurement). Some peak assignments are indicated with the top spectrum.

included in the figures or principal component analyses Table S6.

Fig. 10 shows plots of the percentages measured for the alkyl, methyl, and aromatic/olefinic carbon as a function of travel time from the oil source. Thorn and Aiken (1998) reported that the light, paraffinic crude oil that spilled at the Bemidji site is composed of 87% aliphatic carbons by NMR spectroscopy (Thorn and Aiken, 1998). Therefore, it is reasonable that the  $\text{DOM}_{\text{HC}}$  in the well adjacent to that oil source is composed of 66% alkyl (including 35% methyl) carbons. Across the gradient, the percentage of alkyl carbons decreases to 54% (Fig. 10a, Table S6). The methyl carbons in the DOM decrease by 57% to a final value of 15% (Fig. 10b). These changes indicate that the relatively aliphatic  $\text{DOM}_{\text{HC}}$  is bioavailable and degraded by microbial



**Fig. 10.** Changes in the percentages of (a) alkyl, (b) methyl, and (c) aromatic/olefinic compounds in the pool of  $\text{DOM}_{\text{HC}}$  as a function of time from the center of the oil body. The gray lines indicate the values measured for DOM collected from background well 310E.

communities (Hur et al., 2011). Moreover, aromatic/olefinic compounds comprise a minimum of 9.5% in the well nearest the oil source and there is an increase downgradient in the percentage of aromatics to 17% (Fig. 10c). These results show the selective degradation of highly alkylated, methylated, aliphatic, reduced compounds and preservation of unsaturated, highly oxygenated compounds.

Fig. S1 and Table S6 show the trends for the oxygen-containing functionalities measured by NMR spectroscopy across the plume transect. Ketones comprise 2% of the DOM pool in the wells directly adjacent to the oil source. As the  $\text{DOM}_{\text{HC}}$  pool decreases, the percentage of ketones in the mixture remains constant 188 m (8.6 y) downgradient before increasing to 2.5% in well 925D, located 254 m (11.6 y) downgradient. The percentage of COO/amide compounds is 10.5% in the well adjacent to the oil body and increases in the percentage of COO/amide functional groups to 13% in the  $\text{DOM}_{\text{HC}}$  downgradient from the oil source. Conversely, the class of nitrogen-containing aliphatics ( $\text{H}/\text{C} \geq 1.5$ ,  $\text{N} > 0$ ) measured by UHR-MS ranges from 1.0% adjacent to the source and decreases to 0.045% in the 925D well located 254 m (11.6 y) downgradient. Therefore, we propose that the increase in the COO/amide functional groups across the plume transect is a result of an increase of COO as opposed to amides. The (hemi)acetal, (hemi)ketal carbon class of compounds exhibits a minimum of 2% adjacent to the oil body and remains relatively constant, only increasing to a maximum of 3% downgradient from the source. Finally, the alcohol/ether functional groups exhibit a minimum of 10% in the well adjacent to the oil source. The values for this class increase to 12% downgradient. Although the composition of oxygen-containing functionalities identified in the plume  $\text{DOM}_{\text{HC}}$  trend toward that of the composition of the native groundwater DOM, the  $\text{DOM}_{\text{HC}}$  in well 925D still contains a relatively larger percentage of alkyl groups and a smaller percentage of oxygen-containing groups (Table S6). This result indicates that a portion of the downgradient DOM is a product of the aliphatic oil body from which it is derived as opposed to compounds produced by biosynthesis.

Collectively, the trends observed for the structural changes of  $\text{DOM}_{\text{HC}}$  in the plume by NMR spectroscopy complement those measured by optical spectroscopy, UHR-MS (chemical composition), and BPCA method (aromatic cores). In the context of a pool of DOC that is decreasing as a function of time from the source (Fig. 2), selective preservation is shifting the DOM pool from reduced-aliphatic compounds to highly oxygenated biorefractory products that are similar to CRAM-like compounds. Simultaneous decreases in methyl and alkyl functionalities, coupled with decreases in short wavelength fluorescence, and a decrease in high H/C compounds indicate that the relatively aliphatic DOM is biolabile (Harriman et al., 2017; Spencer et al., 2019; Yamashita and Tanoue, 2003; Davis and Benner, 2007; Heslop et al., 2019). Biodegradation of these compounds is sequential and progressive which enables us to observe a shift in the composition of the DOM pool. The products are characterized by red-shifted fluorescence

and molecular formulae with high O/C (unsaturated, CRAM). The NMR data suggest that these oxygen-containing compounds are predominantly comprised of COO and alcohol or ether functionalities.

#### 4. Implications for understanding plume stability and fate of DOM<sub>HC</sub>

##### 4.1. Plume degradation and source of relatively high-MW compounds

As previously mentioned, there are two proposed mechanisms for the formation of the observed OCOCs in groundwater plumes. The biosynthesis hypothesis proposes that OCOCs are products of microbial growth and biosynthesis, in which hydrocarbon biodegradation drives anabolic production of relatively high-MW biomolecules such as proteins, carbohydrates, lipids, and nucleic acids that are released by the microbes as OCOCs (Mohler et al., 2020). Support of the biosynthesis model is based on the identification of high-MW compounds identified downgradient at the Bemidji site (Mohler et al., 2020). However, the exponential decrease in DOC concentration observed in Fig. 2 and previous studies (Baedecker et al., 1993; Ng et al., 2015; Essaid et al., 2003; Cozzarelli et al., 2010) indicates that the compounds nearest to the source are susceptible to biodegradation and the refractory compounds persist downgradient (i.e., the continuum model). Fig. 11 summarizes the changes in chemical composition of DOM<sub>HC</sub> in the plume. The two red lines indicate the composition of the DOM<sub>HC</sub> adjacent to the oil body and that which persists at the toe of the plume, respectively. The direction of flow can be indicated by the blue arrow because the distance/time from the source also corresponds to an increase in PCA 1 (Fig. 11). The green line indicates the composition of the native DOM in the groundwater collected upgradient from the source. Collectively, Figs. 2 and 9 show a relationship between reactivity and the chemical composition of DOM<sub>HC</sub>. The decrease in DOC concentration corresponds with a change in the composition of the pool of DOM<sub>HC</sub> as biodegradation removes the

most labile compounds. After depleting the most labile fraction, the microbes metabolize the relatively oxidized, aromatic compounds at a slower rate. Finally, the remaining DOM<sub>HC</sub> that has traveled for ~11.6 years (assuming 0.06 m d<sup>-1</sup> flow) (Essaid et al., 2003) to well 925D (located 254 m downgradient from the oil source) is comprised of microbially refractory products with structures consistent with CRAM. These relatively high MW compounds are the result of selective biodegradation of the “light” compounds and persistence of the “heavy” compounds (parent crude has 4–6% resins, and 1–2% asphaltenes).

The persistence of DOM<sub>HC</sub> that is sourced from the petroleum body is further supported by compositional discontinuities identified in the data (Fig. 11). Although the chemical composition of the compounds in the plume are trending towards those identified in the native groundwater DOM in well 310E (right-side of Fig. 11), the molecular composition of the DOM<sub>HC</sub> and native DOM in the groundwater do not converge. This result is visualized as a compositional discontinuity in the PCA plot between well 925D located 254 m (11.6 y) downgradient from the oil body and well 310E located upgradient from the oil body (Fig. 11). The presence of the discontinuity is especially noteworthy because the concentration of DOC at the last well (925D) is less than twice that of the background well (310E) and presumably both the DOM<sub>HC</sub> and the native DOM are present in this well. One possible reason for the discontinuity is that more than 11.6 years are required to degrade the petroleum-derived compounds to products with similar composition as the native DOM. Alternatively, it is possible that the compounds that are produced from this organic matter source (i.e., light aliphatic crude oil) do not produce products that fill in the compositional space between refractory products derived from the oil body and native groundwater DOM at the site. Regardless, the discontinuity is interpreted as an indicator of DOM with different chemical compositions derived from different sources. Finally, support for the degradation continuum model over the biosynthesis model is visualized in Fig. 11. Like the parent crude, the DOM<sub>HC</sub> adjacent to the oil body is composed of low MW, highly reduced, aliphatic,

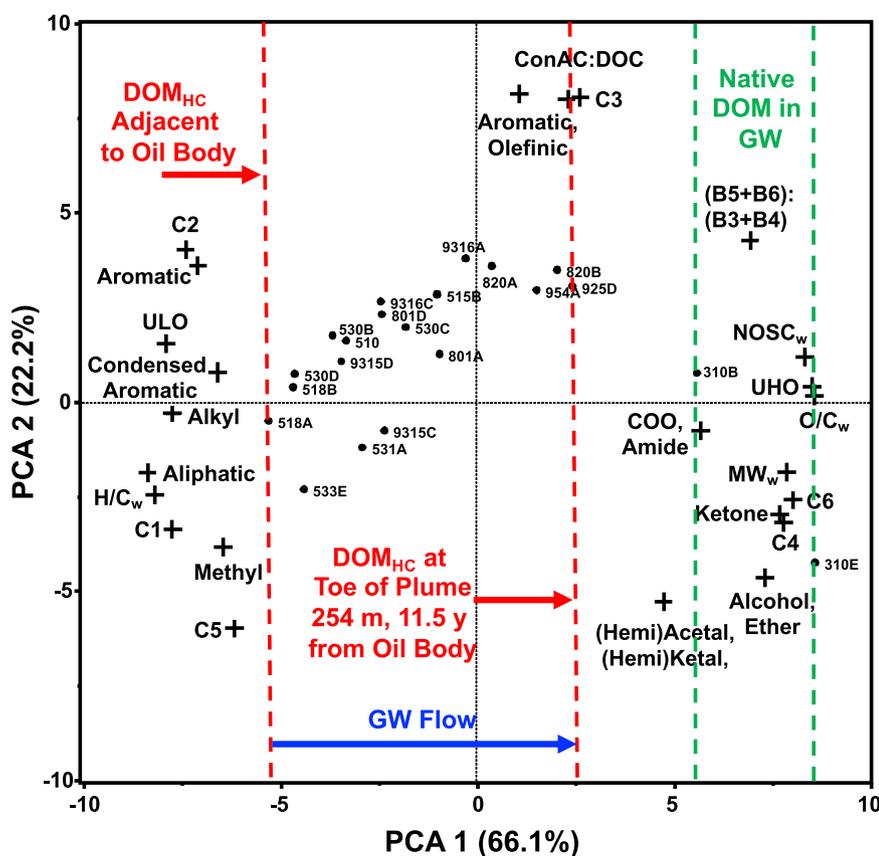


Fig. 11. PCA plot with EEMS, UHR-MS, and NMR data showing the continuum that results from biodegradation of the DOM<sub>HC</sub> from the source to the toe of the plume (right to left). The blue arrow indicates the direction of groundwater flow. Note that the background wells (green line) are upgradient from the oil body and not in the flow path of the contaminated water. Dots indicate groundwater monitoring wells and crosses are PCA loading variables. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and methylated compounds. Selective biodegradation of the most labile components in succession gives rise to the systematic, continuous changes in the composition of the DOM<sub>HC</sub> in the wells between the source and toe of the plume (Fig. 11) (Podgorski et al., 2018).

#### 4.2. Linear vs. non-linear models

We have shown that quantitative measurements (DOC concentration, fluorescence intensity, ConAC concentration, and even the semi-quantitative NMR results to an extent) decrease exponentially while relative changes in DOM<sub>HC</sub> composition change linearly as a function of estimated time traveled from the center of the oil body. A possible explanation for this behavior is that the difference in relative proportions of the compositional endmembers. In this case, the reactive (reduced, aliphatic, low MW) DOM<sub>HC</sub> is at much higher concentration than the background native DOM endmember and degradation products. As such, the change in chemical composition will appear linear until a large fraction of the reactive pool is mineralized or processed. A long linear change is expected along the core of petroleum contaminated plumes followed by a quick, nonlinear change in DOM composition as zero reactive DOM<sub>HC</sub> contamination is approached resulting in a “hockey stick” shape. One way to test this hypothesis would be to sample the toe of the Bemidji plume at higher resolution to determine if there is an asymptote in DOC concentration and corresponding DOM<sub>HC</sub> composition. Unfortunately, the well coverage at the toe of the plume at the Bemidji site is currently limited. However, there are reports of laboratory biodegradation experiments on DOM that demonstrate this concept (Spencer et al., 2009; Sleighter et al., 2014). Sleighter et al. (2014), show exponential decay models of labile, semi-labile, and refractory fractions of DOM from headwater streams that were passed through a bioreactor (Sleighter et al., 2014). All three fractions exhibit degradation patterns that resemble a hockey stick. Most notably is the labile fraction that shows the most rapid linear degradation trend (hockey stick handle) followed by an abrupt change to non-linear change (hockey stick blade) as the labile substrate is consumed. Similarly, it was shown that the early photodegradation timepoints (first 15 days) for DOM from the Congo River exhibit a linear change in chemical composition as highly photolabile DOM is degraded (Spencer et al., 2009). The addition of a 57-day timepoint shows that the change in composition is not linear because it exhibits a quick non-linear change after ~15 days.

The implication for understanding petroleum-derived groundwater plumes is that there is a large pool of highly reduced, aliphatic (in the case of the Bemidji site) carbon that sits on an aquifer that contains a low concentration of relatively biorefractory DOM. Biodegradation of that highly reduced, aliphatic pool of carbon produces partially oxidized intermediates that can partition across the oil/water interface, resulting in DOM<sub>HC</sub>. The proportion of this highly biolabile DOM<sub>HC</sub> is in great excess relative to the background native DOM present in the aquifer. As such, the degradation of these highly biolabile DOM<sub>HC</sub> components initially exhibit linear changes in composition. Measurements of selective fractions of the DOM<sub>HC</sub> pool show that there is a reactivity continuum that spans from labile, semi-labile, and refractory (Sleighter et al., 2014). As shown in Fig. 2 and mentioned previously, more sampling at the toe of the plume is required to identify the presence of asymptote. We posit that there is an asymptote and that even the linear trends that we report will rapidly change to non-linear at that point. Moreover, the composition of the petroleum source is expected to have an impact on the linear and non-linear regions in respect to changes in chemical composition along the core of the plume. This concept will be the subject of future hypotheses that will be tested at spill sites with longer plumes and parent petroleum sources with different chemical compositions.

## 5. Conclusions

The data presented here support a continuum model to explain DOM<sub>HC</sub> dynamics at petroleum-contaminated sites. Although the measurements were made at a single site contaminated with crude oil, the results have generated several hypotheses that can be tested at sites contaminated with refined fuels and crude oils with different compositions. One hypothesis that is of particular interest to regulators is that the DOM<sub>HC</sub> plume biodegradation continuum model suggests that the composition, reactivity, and fate of OCOs are dependent on the composition of the parent petroleum source. As such, follow-up experiments are in progress to examine plume composition originating from petroleum sources with different chemical compositions. If the continuum model is validated from data collected at different sites, the model can be used for predicting plume stability and the composition of compounds that may impact downgradient receptors. While ultimately it may be necessary to determine relationships between the chemical composition of DOM<sub>HC</sub> and impacts on human health as well as the environment (McGuire et al., 2018), the scope of this work is limited to describing the origin, transport, and fate of DOM<sub>HC</sub>.

#### CRediT authorship contribution statement

David C. Podgorskia and Barbara A. Bekins formulated research questions, objectives, experimental design. David C. Podgorski and Phoebe Zito collected samples. David C. Podgorski, Phoebe Zito, Isabelle M. Cozzarelli, Donald F. Smith, Xiaoyan Cao, Klaus Schmidt-Rohr, Sasha Wagner, Aron Stubbins, and Robert G.M. Spencer collected, facilitated, or supported the acquisition of data. Anne M. Kellerman performed statistical analysis. David C. Podgorskia composed the original draft of manuscript and figures with input from all authors. All authors revised the manuscript for publication.

#### Declaration of Competing Interest

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

#### Acknowledgments

This work was supported by Shell Global Solutions (David C. Podgorski, Klaus Schmidt-Rohr), Enbridge Ecofootprint Grant (Project #1703033) (David C. Podgorski, Robert G.M. Spencer), and US Geological Survey Toxic Substances Hydrology Program (Barbara A. Bekins, Isabelle M. Cozzarelli). A portion of this work was performed at the National High Magnetic Field Laboratory ICR User Facility, which is supported by the National Science Foundation Division of Chemistry through DMR-1644779 and the State of Florida. David C. Podgorski and Robert G.M. Spencer were partially supported by NSF OCE-1333157 and OCE-1464396 (Robert G. M. Spencer). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government. The authors thank Jared Trost, Andrew Berg, Jeanne Jaeschke, and the rest of the Bemidji research team for their assistance. The authors thank Jay Brandes for the use of his HPLC instrument for BPCA analysis. The authors are grateful to the late George R. Aiken for his initiation of this work. The authors thank the three anonymous reviewers and Kevin A. Thorn, Emeritus, USGS Earth Surface Processes Division for their contributions.

#### Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.jhazmat.2020.123998](https://doi.org/10.1016/j.jhazmat.2020.123998).

## References

- Aeppli, C., Nelson, R.K., Carmichael, C.A., Arakawa, N., Aluwihare, L.L., Valentine, D.L., Reddy, C.M., 2013. 245th ACS National Meeting, New Orleans LA, Apr 7-11 2013. Characterization of Recalcitrant Oxygenated Hydrocarbons Formed Upon Oil Weathering After the Deepwater Horizon Disaster. American Chemical Society, Washington D.C., p. GEOC-150
- Allan, S.E., Smith, B.W., Anderson, K.A., 2012. Impact of the deepwater horizon oil spill on bioavailable polycyclic aromatic hydrocarbons in Gulf of Mexico coastal waters. *Environ. Sci. Tech.* 46, 2033–2039.
- Altgelt, K.H., Boduszynski, M.M., 1992. Composition of heavy petroleum. 3. An improved boiling-point molecular-weight relation. *Energy Fuels* 6, 68–72.
- Anderson, R.T., Rooney-Varga, J.N., Gaw, C.V., Lovley, D.R., 1998. Anaerobic benzene oxidation in the Fe(III) reduction zone of petroleum contaminated aquifers. *Environ. Sci. Technol.* 32, 1222–1229.
- Baedecker, M.J., Cozzarelli, I.M., Eganhouse, R.P., Siegel, D.I., Bennett, P.C., 1993. Crude-oil in a shallow sand and gravel aquifer. 3. Biogeochemical reactions and mass-balance modeling in anoxic groundwater. *Appl. Geochem.* 8, 569–586.
- Baedecker, M.J., Eganhouse, R.P., Bekins, B.A., Delin, G.N., 2011. Loss of volatile hydrocarbons from an LNAPL oil source. *J. Contam. Hydrol.* 126, 140–152.
- Baedecker, M.J., Eganhouse, R.P., Qi, H., Cozzarelli, I.M., Trost, J.J., Bekins, B.A., 2018. Weathering of oil in a surficial aquifer. *Groundwater* 56, 797–809.
- Baker, A., Lamont-Black, J., 2001. Fluorescence of dissolved organic matter as a natural tracer of ground water. *Groundwater* 39, 745–750.
- Beaver, C.L., Williams, A.E., Atekwana, E.A., Mewafy, F.M., Abdel Aal, G., Slater, L.D., Rossbach, S., 2016. Microbial communities associated with zones of elevated magnetic susceptibility in hydrocarbon-contaminated sediments. *Geomicrobiol. J.* 33, 441–452.
- Bekins, B.A., Cozzarelli, I.M., Erickson, M.L., Steenson, R.A., Thorn, K.A., 2016. Crude oil metabolites in groundwater at two spill sites. *Groundwater* 54, 681–691.
- Bekins, B.A., Cozzarelli, I.M., 2017. Nonvolatile dissolved organic carbon and diesel range organics concentrations measured in 2016 at the USGS crude oil study site near Bemidji, Minnesota, USA. US Geological Survey data release. <https://doi.org/10.5066/F5067CN5733T>.
- Bekins, B.A., Godsy, E.M., Warren, E., 1999. Distribution of microbial physiologic types in an aquifer contaminated by crude oil. *Microb. Ecol.* 37, 263–275.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Ser. B (Methodol.)* 57, 289–300.
- Benner, R., Amon, R.M.W., 2015. The size-reactivity continuum of major bioelements in the ocean. *Ann. Rev. Mar. Sci.* 7, 185–205.
- Bennett, P.C., Siegel, D.E., Baedecker, M.J., Hult, M.F., 1993. Crude-oil in a shallow sand and gravel aquifer. 1. Hydrogeology and inorganic geochemistry. *Appl. Geochem.* 8, 529–549.
- Bianchi, T.S., Osburn, C., Shields, M.R., Yvon-Lewis, S., Young, J., Guo, L., Zhou, Z., 2014. Deepwater horizon oil in Gulf of Mexico waters after 2 years: transformation into the dissolved organic matter pool. *Environ. Sci. Technol.* 48, 9288–9297.
- Boduszynski, M.M., 1987. Composition of heavy petroleum. 1. Molecular-weight, hydrogen deficiency, and heteroatom concentration as a function of atmospheric equivalent boiling-point up to 1400-degrees-F (760-degrees-C). *Energy Fuels* 1, 2–11.
- Boduszynski, M.M., 1988. Composition of heavy petroleum. 2. Molecular characterization. *Energy Fuels* 2, 597–613.
- Boduszynski, M.M., Altgelt, K.H., 1992. Composition of heavy petroleum. 4. Significance of the extended atmospheric equivalent boiling-point (Aebp) scale. *Energy Fuels* 6, 72–76.
- Callaghan, A.V., 2013. Metabolomic investigations of anaerobic hydrocarbon-impacted environments. *Curr. Opin. Biotechnol.* 24, 506–515.
- Cao, X.Y., Aiken, G.R., Butler, K.D., Huntington, T.G., Balch, W.M., Mao, J.D., Schmidt-Rohr, K., 2018. Evidence for major input of riverine organic matter into the ocean. *Org. Geochem.* 116, 62–76.
- Cao, F., Medeiros, P.M., Miller, W.L., 2016. Optical characterization of dissolved organic matter in the Amazon river plume and the adjacent ocean: examining the relative role of mixing, photochemistry, and microbial alterations. *Mar. Chem.* 186, 178–188.
- Cao, X., 2014. Spectroscopic characterization of dissolved organic matter: insights into the linkage between sources and chemical composition (Chemistry and Biochemistry theses). Old Dominion University. [https://digitalcommons.odu.edu/chemistry\\_etd\\_s/26](https://digitalcommons.odu.edu/chemistry_etd_s/26).
- Catalán, N., Casas-Ruiz, J.P., von Schiller, D., Proia, L., Obrador, B., Zwiemann, E., Marcé, R., 2017. Biodegradation kinetics of dissolved organic matter chromatographic fractions in an intermittent river. *J. Geophys. Res. Biogeosci.* 122, 131–144.
- Chang, Z., Tian, L., Li, F., Zhou, Y., Wu, M., Steinberg, C.E.W., Dong, X., Pan, B., Xing, B., 2018. Benzene polycarboxylic acid – a useful marker for condensed organic matter, but not for only pyrogenic black carbon. *Sci. Total Environ.* 626, 660–667.
- Corilo, Y., 2015. *EnviroOrg*. Florida State University, Tallahassee.
- Cozzarelli, I.M., Baedecker, M.J., Eganhouse, R.P., Goerlitz, D.F., 1994. The geochemical evolution of low-molecular-weight organic acids derived from the degradation of petroleum contaminants in groundwater. *Geochim. Cosmochim. Acta* 58, 863–877.
- Cozzarelli, I.M., Bekins, B.A., Eganhouse, R.P., Warren, E., Essaid, H.I., 2010. In situ measurements of volatile aromatic hydrocarbon biodegradation rates in groundwater. *J. Contam. Hydrol.* 111, 48–64.
- Cozzarelli, I.M., Schreiber, M.E., Erickson, M.L., Ziegler, B.A., 2016. Arsenic cycling in hydrocarbon plumes: secondary effects of natural attenuation. *Groundwater* 54, 35–45.
- D'Andrilli, J., Cooper, W.T., Foreman, C.M., Marshall, A.G., 2015. An ultrahigh-resolution mass spectrometry index to estimate natural organic matter lability. *Rapid Commun. Mass Spectrom.* 29, 2385–2401.
- Davis, J., Benner, R., 2007. Quantitative estimates of labile and semi-labile dissolved organic carbon in the western Arctic Ocean: a molecular approach. *Limnol. Oceanogr.* 52, 2434–2444.
- Dittmar, T., 2008. The molecular level determination of black carbon in marine dissolved organic matter. *Org. Geochem.* 39, 396–407.
- Dittmar, T., Koch, B., Hertkorn, N., Kattner, G., 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr. Methods* 6, 230–235.
- Drake, T.W., Guillemette, F., Hemingway, J.D., Chanton, J.P., Podgorski, D.C., Zimov, N.S., Spencer, R.G.M., 2018. The ephemeral signature of permafrost carbon in an Arctic Fluvial Network. *J. Geophys. Res. Biogeosci.* 123, 1475–1485.
- Dvorski, S.E.M., Gonsior, M., Hertkorn, N., Uhl, J., Muller, H., Griebler, C., Schmitt-Kopplin, P., 2016. Geochemistry of dissolved organic matter in a spatially highly resolved groundwater petroleum hydrocarbon plume cross-section. *Environ. Sci. Technol.* 50, 5536–5546.
- Eganhouse, R.P., Baedecker, M.J., Cozzarelli, I.M., Aiken, G.R., Thorn, K.A., Dorsey, T.F., 1993. Crude-oil in a shallow sand and gravel aquifer. 2. Organic geochemistry. *Appl. Geochem.* 8, 551–567.
- Eganhouse, R.P., Dorsey, T.F., Phinney, C.S., Westcott, A.M., 1996. Processes Affecting the Fate of monoaromatic hydrocarbons in an aquifer contaminated by crude oil. *Environ. Sci. Technol.* 30, 3304–3312.
- EPA, 1999. Use of Monitored Natural Attenuation at Superfund, RCRA Corrective Action, and Underground Storage Tank Sites. EPA, Office of Solid Waste and Emergency Response, Washington, DC.
- Essaid, H.I., Bekins, B.A., Herkelrath, W.N., Delin, G.N., 2011. Crude oil at the Bemidji site: 25 years of monitoring, modeling, and understanding. *Groundwater* 49, 706–726.
- Essaid, H.I., Cozzarelli, I.M., Eganhouse, R.P., Herkelrath, W.N., Bekins, B.A., Delin, G.N., 2003. Inverse modeling of BTEX dissolution and biodegradation at the Bemidji, MN crude-oil spill site. *J. Contam. Hydrol.* 67, 269–299.
- Fahrenfeld, N., Cozzarelli, I.M., Bailey, Z., Pruden, A., 2014. Insights into biodegradation through depth-resolved microbial community functional and structural profiling of a crude-oil contaminant plume. *Microb. Ecol.* 68, 453–462.
- Farrington, J.W., Quinn, J.G., 2015. “Unresolved complex mixture” (UCM): a brief history of the term and moving beyond it. *Mar. Pollut. Bull.* 96, 29–31.
- Fellman, J.B., Hood, E., Spencer, R.G.M., 2010. Fluorescence spectroscopy opens new windows into dissolved organic matter dynamics in freshwater ecosystems: a review. *Limnol. Oceanogr.* 55, 2452–2462.
- Frontera-Suau, R., Bost, F.D., McDonald, T.J., Morris, P.J., 2002. Aerobic biodegradation of hopanes and other biomarkers by crude oil-degrading enrichment cultures. *Environ. Sci. Technol.* 36, 4585–4592.
- Fry, N., Steenson, R.A., 2019. User's guide: derivation and application of environmental screening levels (ESLs). In: Chapter 4 – Methods: Petroleum Mixture Screening Levels. San Francisco Bay Regional Water Quality Control Board, San Francisco, pp. 1–35.
- Fuchs, G., Boll, M., Heider, J., 2011. Microbial degradation of aromatic compounds – from one strategy to four. *Nat. Rev. Microbiol.* 9, 803–816.
- Gieg, L.M., Sufliata, J.M., 2002. Detection of anaerobic metabolites of saturated and aromatic hydrocarbons in petroleum-contaminated aquifers. *Environ. Sci. Technol.* 36, 3755–3762.
- Gieg, L.M., Sufliata, J.M., 2005. Metabolic indicators of anaerobic hydrocarbon biodegradation in petroleum-laden environments. In: Ollivier, B., Magot, M. (Eds.), *Petroleum Microbiology*. American Society of Microbiology, pp. 337–356.
- Guillemette, F., McCallister, S.L., del Giorgio, P.A., 2013. Differentiating the degradation dynamics of algal and terrestrial carbon within complex natural dissolved organic carbon in temperate lakes. *J. Geophys. Res. Biogeosci.* 118, 963–973.
- Harriman, B.H., Zito, P., Podgorski, D.C., Tarr, M.A., Sufliata, J.M., 2017. Impact of photooxidation and biodegradation on the fate of oil spilled during the deepwater horizon incident: advanced stages of weathering. *Environ. Sci. Technol.* 51, 7412–7421.
- Harshman, R., 1984. “How can I know if it's real?” A catalog of diagnostics for use with three-mode factor analysis and multidimensional scaling. In: Law, H.G., Snyder, C. W., Hattie, J.A., McDonald, R.P. (Eds.), *Research Methods for Multimode Data Analysis*. Praeger Publishers Inc, Westport, Connecticut, pp. 566–591.
- Hawkes, J.A., D'Andrilli, J., Sleighter, R.L., Chen, H., Hatcher, P.G., Ijaz, A., Khaksan, M., Schum, S., Mazzoleni, L., Chu, R., Tolic, N., Kew, W., Hess, N., Lv, J., Zhang, S., Chen, H., Shi, Q., Hutchins, R.H.S., Lozano, D.C.P., Gavard, R., Jones, H.E., Thomas, M.J., Barrow, M., Osterholz, H., Dittmar, T., Simon, C., Gleixner, G., Berg, S.M., Remualdo, C.K., Catalán, N., Cole, R.B., Noreiga-Ortega, B., Singer, G., Radoman, N., Schmitt, N.D., Stubbins, A., Agar, J.N., Zito, P., Podgorski, D.C., 2020. An international laboratory comparison of dissolved organic matter composition by high resolution mass spectrometry: are we getting the same answer? *Limnol. Oceanogr. Methods*.
- Hazen, T.C., Prince, R.C., Mahmoudi, N., 2016. Marine oil biodegradation. *Environ. Sci. Technol.* 50, 2121–2129.
- Heider, J., 2007. Adding handles to unhandy substrates: anaerobic hydrocarbon activation mechanisms. *Curr. Opin. Chem. Biol.* 11, 188–194.
- Hemingway, J.D., 2017. Fouriertransform: open-source tools for FT-ICR MS data analysis.
- Hertkorn, N., Frommberger, M., Witt, M., Koch, B.P., Schmitt-Kopplin, P., Perdue, E.M., 2008. Natural organic matter and the event horizon of mass spectrometry. *Anal. Chem.* 80, 8908–8919.

- Heslop, J.K., Winkel, M., Walter Anthony, K.M., Spencer, R.G.M., Podgorski, D.C., Zito, P., Kholodov, A., Zhang, M., Liebner, S., 2019. Increasing organic carbon biolability with depth in Yedoma Permafrost: ramifications for future climate change. *J. Geophys. Res. Biogeosci.* 124, 2021–2038.
- He, C., Zhang, Y., Li, Y., Zhuo, X., Li, Y., Zhang, C., Shi, Q., 2020. In-house standard method for molecular characterization of dissolved organic matter by FT-ICR mass spectrometry. *ACS Omega* 5, 11730–11736.
- Hindersmann, B., Achten, C., 2017. Accelerated benzene polycarboxylic acid analysis by liquid chromatography-time-of-flight-mass spectrometry for the determination of petrogenic and pyrogenic carbon. *J. Chromatogr. A* 1510, 57–65.
- Hur, J., Lee, B.-M., Shin, H.-S., 2011. Microbial degradation of dissolved organic matter (DOM) and its influence on phenanthrene-DOM interactions. *Chemosphere* 85, 1360–1367.
- Islam, A., Ahmed, A., Hur, M., Thorn, K.A., Kim, S., 2016. Molecular-level evidence provided by ultrahigh resolution mass spectrometry for oil-derived DOC in groundwater at Bemidji, Minnesota. *J. Hazard. Mater.* 320.
- Jaggi, A., Radović, J.R., Snowdon, L.R., Larter, S.R., Oldenburg, T.B.P., 2019. Composition of the dissolved organic matter produced during in situ burning of spilled oil. *Org. Geochem.* 138, 103926.
- Jobelius, C., Ruth, B., Griebler, C., Meckenstock, R.U., Hollender, J., Reineke, A., Frimmel, F.H., Zwiener, C., 2011. Metabolites indicate hot spots of biodegradation and biogeochemical gradients in a high-resolution monitoring well. *Environ. Sci. Technol.* 45, 474–481.
- Johnson, R.L., Schmidt-Rohr, K., 2014. Quantitative solid-state  $^{13}\text{C}$  NMR with signal enhancement by multiple cross polarization. *J. Magn. Reson.* 239, 44–49.
- Kanally, R.A., Harayama, S., 2000. Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by bacteria. *J. Bacteriol.* 182, 2059–2067.
- Kellerman, A.M., Guillemette, F., Podgorski, D.C., Aiken, G.R., Butler, K.D., Spencer, R.G.M., 2018. Unifying concepts linking dissolved organic matter composition to persistence in aquatic ecosystems. *Environ. Sci. Technol.* 52, 2538–2548.
- Kellerman, A.M., Kothawala, D.N., Dittmar, T., Tranvik, L.J., 2015. Persistence of dissolved organic matter in lakes related to its molecular characteristics. *Nat. Geosci.* 8, 454–457.
- Kim, S., Kramer, R.W., Hatcher, P.G., 2003. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the Van Krevelen diagram. *Anal. Chem.* 75, 5336–5344.
- Koch, B.P., Dittmar, T., 2006. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* 20, 926–932.
- Koch, B.P., Witt, M., Engbrodt, R., Dittmar, T., Kattner, G., 2005. Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochim. Cosmochim. Acta* 69, 3299–3308.
- Koenig, J., Balfanz, E., Funcke, W., Romanowski, T., 1983. Determination of oxygenated polycyclic aromatic hydrocarbons in airborne particulate matter by capillary gas chromatography and gas chromatography/mass spectrometry. *Anal. Chem.* 55, 599–603.
- Kowalczyk, P., Cooper, W.J., Whitehead, R.F., Durako, M.J., Sheldon, W., 2003. Characterization of CDOM in an organic-rich river and surrounding coastal ocean in the South Atlantic Bight. *Aquat. Sci.* 65, 384–401.
- Kulkarni, H., Mladenov, N., Datta, S., 2019. Effects of acidification on the optical properties of dissolved organic matter from high and low arsenic groundwater and surface water. *Sci. Total Environ.* 653, 1326–1332.
- Kuzyakov, Y., Bogomolova, I., Glaser, B., 2014. Biochar stability in soil: decomposition during eight years and transformation as assessed by compound-specific  $^{14}\text{C}$  analysis. *Soil Biol. Biochem.* 70, 229–236.
- Lechtenfeld, O.J., Kattner, G., Flerus, R., McCallister, S.L., Schmitt-Kopplin, P., Koch, B.P., 2014. Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean. *Geochim. Cosmochim. Acta* 126, 321–337.
- Lundstedt, S., White, P.A., Lemieux, C.L., Lynes, K.D., Lambert, L.B., Oberg, L., Haglund, P., Tysklind, M., 2007. Sources, fate, and toxic hazards of oxygenated polycyclic aromatic hydrocarbons (PAHs) at PAH-contaminated sites. *AMBIO* 36, 475–485.
- Luzius, C., Guillemette, F., Podgorski, D.C., Kellerman, A.M., Spencer, R.G.M., 2018. Drivers of dissolved organic matter in the vent and major conduits of the world's largest freshwater spring. *J. Geophys. Res. Biogeosci.* 123, 2775–2790.
- Mao, J.-D., Schmidt-Rohr, K., 2004. Accurate quantification of aromaticity and nonprotonated aromatic carbon fraction in natural organic matter by  $^{13}\text{C}$  solid state nuclear magnetic resonance. *Environ. Sci. Technol.* 38, 2680–2684.
- McGuire, J.T., Cozzarelli, I.M., Bekins, B.A., Link, H., Martinovic-Weigelt, D., 2018. Toxicity assessment of groundwater contaminated by petroleum hydrocarbons at a well-characterized, aged, crude oil release site. *Environ. Sci. Technol.* 52, 12172–12178.
- Meador, J.P., Nahrgang, J., 2019. Characterizing crude oil toxicity to early-life stage fish based on a complex mixture: are we making unsupported assumptions? *Environ. Sci. Technol.* 53, 11080–11092.
- Mohler, R.E., Ahn, S., O'Reilly, K., Zemo, D.A., Espino Devine, C., Magaw, R., Sihota, N., 2020. Towards comprehensive analysis of oxygen containing organic compounds in groundwater at a crude oil spill site using GC×GC-TOFMS and Orbitrap ESI-MS. *Chemosphere* 244, 125504.
- Mohler, R.E., O'Reilly, K.T., Zemo, D.A., Tiwary, A.K., Magaw, R.I., Synowicz, K.A., 2013. Non-targeted analysis of petroleum metabolites in groundwater using GC×GC-TOFMS. *Environ. Sci. Technol.* 47, 10471–10476.
- Mostovaya, A., Hawkes, J.A., Dittmar, T., Tranvik, L.J., 2017. Molecular determinants of dissolved organic matter reactivity in lake water. *Front. Earth Sci.* 5, 1–13.
- Mostovaya, A., Hawkes, J.A., Koehler, B., Dittmar, T., Tranvik, L.J., 2017. Emergence of the reactivity continuum of organic matter from kinetics of a multitude of individual molecular constituents. *Environ. Sci. Technol.* 51, 11571–11579.
- Murphy, K.R., Ruiz, G.M., Dunsmuir, W.T.M., Waite, T.D., 2006. Optimized parameters for fluorescence-based verification of ballast water exchange by ships. *Environ. Sci. Technol.* 40, 2357–2362.
- Murphy, K.R., Stedmon, C.A., Graeber, D., Bro, R., 2013. Fluorescence spectroscopy and multi-way techniques. *PARAFAC. Anal. Methods* 5, 6557–6566.
- Murphy, K.R., Stedmon, C.A., Wenig, P., Bro, R., 2014. OpenFluor— an online spectral library of auto-fluorescence by organic compounds in the environment. *Anal. Methods* 6, 658–661.
- Ng, G.H.C., Bekins, B.A., Cozzarelli, I.M., Baedecker, M.J., Bennett, P.C., Amos, R.T., Herkelrath, W.N., 2015. Reactive transport modeling of geochemical controls on secondary water quality impacts at a crude oil spill site near Bemidji, MN. *Water Resour. Res.* 51, 4156–4183.
- Nicol, S., Dugay, J., Hennion, M.C., 2001. Determination of oxygenated polycyclic aromatic compounds in airborne particulate organic matter using gas chromatography-tandem mass spectrometry. *Chromatographia* 53, S464–S469.
- NRC (Ed.), 2012. In: Cleanup of Most Challenging U.S. Contaminated Groundwater Sites Unlikely for Many Decades. National Academies of Sciences, Engineering, and Medicine, Washington.
- O'Donnell, J.A., Aiken, G.R., Butler, K.D., Guillemette, F., Podgorski, D.C., Spencer, R.G.M., 2016. DOM composition and transformation in boreal forest soils: the effects of temperature and organic-horizon decomposition state. *J. Geophys. Res. Biogeosci.* 121, 2727–2744.
- O'Reilly, K.T., 2018. Comment on “examining natural attenuation and acute toxicity of petroleum-derived dissolved organic matter with optical spectroscopy”. *Environ. Sci. Technol.* 52, 11960–11961.
- O'Reilly, K.T., Mohler, R.E., Zemo, D.A., Ahn, S., Tiwary, A.K., Magaw, R.I., Espino, D.C., Synowicz, K.A., 2015. Identification of ester metabolites from petroleum hydrocarbon biodegradation in groundwater using GC×GC-TOFMS. *Environ. Toxicol. Chem.* 34, 1959–1961.
- O'Reilly, K.T., Mohler, R.E., Zemo, D.A., Ahn, S., Magaw, R.I., Espino Devine, C., 2019. Oxygen-containing compounds identified in groundwater from fuel release sites using GC×GC-TOF-MS. *Groundw. Monit. Remediat.* 39, 32–40.
- Oberling, L.K., Gieg, L.M., 2018. Methanogenic paraffin biodegradation: alkylsuccinate synthase gene quantification and dicarboxylic acid production. *Appl. Environ. Microbiol.* 84, e01773-01717.
- Ohno, T., 2002. Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. *Environ. Sci. Technol.* 36, 742–746.
- Osburn, C., Boyd, T.J., Montgomery, M.T., Bianchi, T.S., Coffin, R.B., W, P.H., 2016. Optical proxies for terrestrial dissolved organic matter in estuaries and coastal waters. *Front. Mar. Sci.* 2, 1–15.
- Overton, E.B., Wade, T.L., Radović, J.R., Meyer, B.M., Miles, M.S., Larter, S.R., 2016. Chemical composition of Macondo and other crude oils and compositional alterations during oil spills. *Oceanography* 29, 50–63.
- Podgorski, D.C., Zito, P., McGuire, J.T., Martinovic-Weigelt, D., Cozzarelli, I.M., Bekins, B.A., Spencer, R.G.M., 2018. Examining natural attenuation and acute toxicity of petroleum-derived dissolved organic matter with optical spectroscopy. *Environ. Sci. Technol.* 52, 6157–6166.
- , 2015RC Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Riedel, T., Biester, H., Dittmar, T., 2012. Molecular fractionation of dissolved organic matter with metal salts. *Environ. Sci. Technol.* 46, 4419–4426.
- Roebuck, J.A., Seidel, M., Dittmar, T., Jaffé, R., 2018. Land use controls on the spatial variability of dissolved black carbon in a subtropical watershed. *Environ. Sci. Technol.* 52, 8104–8114.
- Šantl-Temkiv, T., Finster, K., Dittmar, T., Hansen, B.M., Thyraug, R., Nielsen, N.W., Karlson, U.G., 2013. Hailstones: a window into the microbial and chemical inventory of a storm cloud. *PLOS One* 8, e53550.
- Sauer, T.C., Brown, J.S., Boehm, P.D., Aurand, D.V., Michel, J., Hayes, M.O., 1993. Hydrocarbon source identification and weathering characterization of intertidal and subtidal sediments along the Saudi Arabian coast after the Gulf War oil spill. *Mar. Pollut. Bull.* 27, 117–134.
- Schlau, J.E., Kramer, A.L., Chlebowska, A., Truong, L., Tanguay, R.L., Simonich, S.L.M., Semprini, L., 2017. Formation of developmentally toxic phenanthrene metabolite mixtures by *Mycobacterium* sp. ELW1. *Environ. Sci. Technol.* 51, 8569–8578.
- Simarro, R., González, N., Bautista, L.F., Molina, M.C., 2013. Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by a wood-degrading consortium at low temperatures. *FEMS Microbiol. Ecol.* 83, 438–449.
- Sivaram, A.K., Logeshwaran, P., Lockington, R., Naidu, R., Megharaj, M., 2019. Low molecular weight organic acids enhance the high molecular weight polycyclic aromatic hydrocarbons degradation by bacteria. *Chemosphere* 222, 132–140.
- Sleighter, R.L., Cory, R.M., Kaplan, L.A., Abdulla, H.A.N., Hatcher, P.G., 2014. A coupled geochemical and biogeochemical approach to characterize the bioreactivity of dissolved organic matter from a headwater stream. *J. Geophys. Res. Biogeosci.* 119, 1520–1537.
- Smith, D.F., Podgorski, D.C., Rodgers, R.P., Blakney, G.T., Hendrickson, C.L., 2018. 21 Tesla FT-ICR mass spectrometer for ultrahigh-resolution analysis of complex organic mixtures. *Anal. Chem.* 90, 2041–2047.
- Spencer, R.G.M., Butler, K.D., Aiken, G.R., 2012. Dissolved organic carbon and chromophoric dissolved organic matter properties of rivers in the USA. *J. Geophys. Res. Biogeosci.* 117, 1–14.
- Spencer, R.G.M., Guo, W., Raymond, P.A., Dittmar, T., Hood, E., Fellman, J., Stubbins, A., 2014. Source and biolability of ancient dissolved organic matter in

- glacier and lake ecosystems on the Tibetan Plateau. *Geochim. Cosmochim. Acta* 142, 64–74.
- Spencer, R.G.M., Kellerman, A.M., Podgorski, D.C., Macedo, M.N., Jankowski, K., Nunes, D., Neill, C., 2019. Identifying the molecular signatures of agricultural expansion in amazonian headwater streams. *J. Geophys. Res. Biogeosci.* 124, 1637–1650.
- Spencer, R.G.M., Stubbins, A., Hernes, P.J., Baker, A., Mopper, K., Aufdenkampe, A.K., Dyda, R.Y., Mwamba, V.L., Mangangu, A.M., Wabakanghanzi, J.N., Six, J., 2009. Photochemical degradation of dissolved organic matter and dissolved lignin phenols from the Congo River. *J. Geophys. Res. Biogeosci.* 114.
- Stedmon, Ca, Bro, R., 2008. Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. *Limnol. Oceanogr. Methods* 6, 572–579.
- Stubbins, A., Spencer, R.G.M., Chen, H., Hatcher, P.G., Mopper, K., Hernes, P.J., Mwamba, V.L., Mangangu, A.M., Wabakanghanzi, J.N., Six, J., 2010. Illuminated darkness: molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry. *Limnol. Oceanogr.* 55, 1467–1477.
- Stubbins, A., Spencer, R.G.M., Mann, P.J., Holmes, R.M., McClelland, J.W., Niggemann, J., Dittmar, T., 2015. Utilizing colored dissolved organic matter to derive dissolved black carbon export by arctic rivers. *Front. Earth Sci.* 3.
- Tfaily, M.M., Podgorski, D.C., Corbett, J.E., Chanton, J.P., Cooper, W.T., 2011. Influence of acidification on the optical properties and molecular composition of dissolved organic matter. *Anal. Chim. Acta* 706, 261–267.
- Thorn, K.A., Aiken, G.R., 1998. Biodegradation of crude oil into nonvolatile organic acids in a contaminated aquifer near Bemidji, Minnesota. *Org. Geochem.* 29, 909–931.
- Tomco, P.L., Zulueta, R.C., Miller, L.C., Zito, P., Campbell, R.W., Welker, J.M., 2019. DOC export is exceeded by C fixation in May Creek: a late-successional watershed of the Copper River Basin, Alaska. *PLOS One* 14, e0225271.
- Toth, C.R.A., Gieg, L.M., 2018. Time course-dependent methanogenic crude oil biodegradation: dynamics of fumarate addition metabolites, biodegradative genes, and microbial community composition. *Front. Microbiol.* 8, 2610.
- Vähätalo, A.V., Aarnos, H., Mäntyniemi, S., 2010. Biodegradability continuum and biodegradation kinetics of natural organic matter described by the beta distribution. *Biogeochemistry* 100, 227–240.
- Del Vecchio, R., Blough, N.V., 2004. Spatial and seasonal distribution of chromophoric dissolved organic matter and dissolved organic carbon in the Middle Atlantic Bight. *Mar. Chem.* 89, 169–187.
- Wagner, S., Brandes, J., Goranov, A.I., Drake, T.W., Spencer, R.G.M., Stubbins, A., 2017. Online quantification and compound-specific stable isotopic analysis of black carbon in environmental matrices via liquid chromatography-isotope ratio mass spectrometry. *Limnol. Oceanogr. Methods* 15, 995–1006.
- Wagner, S., Ding, Y., Jaffé, R., 2017. A new perspective on the apparent solubility of dissolved black carbon. *Front. Earth Sci.* 5.
- Yamashita, Y., Tanoue, E., 2003. Distribution and alteration of amino acids in bulk DOM along a transect from bay to oceanic waters. *Mar. Chem.* 82, 145–160.
- Yan, M., Fu, Q., Li, D., Gao, G., Wang, D., 2013. Study of the pH influence on the optical properties of dissolved organic matter using fluorescence excitation-emission matrix and parallel factor analysis. *J. Lumin.* 142, 103–109.
- Zemo, D.A., Foote, G.R., 2003. The technical case for eliminating the use of the TPH analysis in assessing and regulating dissolved petroleum hydrocarbons in ground water. *Groundw. Monit. Remediat.* 23, 95–104.
- Zemo, D.A., O'Reilly, K.T., Mohler, R.E., Magaw, R.I., Espino Devine, C., Ahn, S., Tiwary, A.K., 2017. Life cycle of petroleum biodegradation metabolite plumes, and implications for risk management at fuel release sites. *Integr. Environ. Assess. Manag.* 13, 714–727.
- Zemo, D.A., O'Reilly, K.T., Mohler, R.E., Tiwary, A.K., Magaw, R.I., Synowiec, K.A., 2013. Nature and estimated human toxicity of polar metabolite mixtures in groundwater quantified as TPHd/DRO at biodegrading fuel release sites. *Groundw. Monit. Remediat.* 33, 44–56.
- Zito, P., Ghannam, R., Bekins, B.A., Podgorski, D.C., 2019. Examining the extraction efficiency of petroleum-derived dissolved organic matter in contaminated groundwater plumes. *Groundw. Monit. Remediat.* 39, 25–31.
- Zito, P., Podgorski, D.C., Johnson, J., Chen, H., Rodgers, R.P., Guillemette, F., Kellerman, A.M., Spencer, R.G.M., Tarr, M.A., 2019. Molecular-level composition and acute toxicity of photosolubilized petrogenic carbon. *Environ. Sci. Technol.* 53, 8235–8243.