Supporting Information

Geochemistry of Dissolved Organic Matter in a Spatially Highly Resolved Groundwater Petroleum Hydrocarbon Plume Cross-Section

*Sabine E.-M. Dvorski [a], Michael Gonsior [b], Norbert Hertkorn [a], Jenny Uhl [a], Hubert Müller [c], Christian Griebler [c], Philippe Schmitt-Kopplin** [a,d]

[a] Helmholtz Zentrum München - German Research Center for Environmental Health, Research Unit Analytical BioGeoChemistry, D-85764 Neuherberg, Germany

[b] University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory, Solomons, Maryland 20688, United States

[c] Helmholtz Zentrum München - German Research Center for Environmental Health, Institute of Groundwater Ecology, D-85764 Neuherberg, Germany

[d] Technische Universität München, Chair of Analytical Food Chemistry, D-85354 Freising-Weihenstephan, Germany

* Corresponding author: Philippe Schmitt-Kopplin, Helmholtz Zentrum München, - German Research Center for Environmental Health, Research Unit Analytical BioGeoChemistry, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany, phone: (+49) 089 3187 3246, fax: (+49) 089 3187 3358, Email: schmitt-kopplin@helmholtz-muenchen.de

This supporting information contains 21 pages including 8 figures and 5 tables.

Contents

Methods

Sampling depths

Thirty-two groundwater samples were taken along the high-resolution multilevel well (HR-MLW) at the following depths. Samples marked with * were taken in duplicates with a sample volume of 250 mL and 1 L, respectively.

6.51 m bls, 6.54 m bls*, 6.59 m bls, 6.61 m bls, 6.64 m bls, 6.67 m bls*, 6.75 m bls, 6.78 m bls*, 6.81 m bls, 6.83 m bls, 6.88 m bls, 6.91 m bls, 6.93 m bls, 6.96 m bls, 7.01 m bls, 7.03 m bls, 7.06 m bls*, 7.08 m bls, 7.11 m bls, 7.21 m bls*, 7.26 m bls, 7.31 m bls, 7.41 m bls, 7.51 m bls, 7.61 m bls*, 7.75 m bls, 8.05 m bls*, 8.65 m bls, 9.05 m bls*, 9.35 m bls, 10.2 m bls, 11.19 m bls.

Solid phase extraction of DOM

DOM from the 0.2 µm filtered 250 mL and 1 L groundwater samples was isolated by an established solid phase extraction (SPE) method described by Dittmar et al.[1] Briefly, the water samples (250 mL and 1 L) were acidified to pH 2 with hydrochloric acid (32%, p.a., Merck KGaA, Darmstadt, Germany) and passed through Agilent Bond Elut PPL SPE cartridges (100 mg and 500 mg with a flowrate of <4 mL min^{-1} and <10 mL min^{-1}, respectively). To prevent loss of organics caused by overloading of the first SPE cartridge, a second cartridge was attached below the first one. Then, the cartridges were rinsed with acidified (pH 2) purified water (MilliQ-Integral, Merck KGaA, Darmstadt, Germany) and dried under vacuum. The DOM was eluted with 2 mL (100 mg PPL cartridges) and 10 mL (500 mg PPL cartridges) methanol (Chromasolv® LC-MS grade methanol, Sigma Aldrich, Taufkirchen, Germany), respectively.

Ultrahigh resolution mass spectrometry

Ultrahigh resolution mass spectra were acquired with a Bruker (Bremen, Germany) SolariX Qe FT-ICR-MS equipped with a 12 T superconducting magnet. Prior to (-)ESI FT-ICR-MS analysis, the SPE-DOM

samples were diluted 1:20 with methanol and continuously infused with a flowrate of 120 µL h^{-1}. The dilution of the samples ensured a good compromise of signal intensity and comparable transient spectra in the ion cyclotron resonance (ICR) cell (same quantities of molecules entering the ICR cell), as well as avoided effects from ion-ion and ion-molecule interaction (e.g. peak splitting). Spectra were acquired in the range of m/z 123–1000 with 500 scans per spectrum averaged. For (+)APPI, the SPE-DOM extracts were diluted 1:20 with a solution of 90% methanol and 10% toluene as a dopant.[2] The samples were continuously infused with a flowrate of 1 mL h^{-1}. Because of the low molecular mass of typical contamination molecules, spectra were acquired in the range of m/z 100–1000 with 500 scans per spectrum averaged. To prevent carry over and cross-contamination in (+)APPI FT-ICR-MS, the APPI source was rinsed thoroughly with a solution of 90% methanol and 10% toluene between the measurements. Blank methanol samples were frequently recorded and did not show any signs of sample carryover. Additionally, SPE extraction blanks were measured and the sample spectra were corrected after acquisition in Excel by filtering peaks resulting from the SPE resin and always present contaminants from methanol and the instrumentation itself. Reproducibility of the mass spectra was assured by measuring SPE-DOM from 250 mL and 1 L samples collected at the same depths, also indicating no extraction volume effect**.** Additionally, FT-ICR mass spectra of samples taken in close vicinity to each other (enabled by the high spatial resolution sampling in the centimeter range) resemble in principle composition, indicating indirectly reproducibility of FT-ICR mass spectra.

The mass spectra were internally calibrated to known and high abundant masses of DOM with a mass accuracy of 0.1 ppm. For both (-)ESI and (+)APPI FT-ICR-MS spectra, only singly charged molecular ions were found. Peak tables were exported with a signal to noise ratio of ≥ 4.The formula assignment of the mass spectra was performed with the in-house written FormCalc software tool and the NetCalc network approach described previously.[3] The mass accuracy window for the formula assignment was set to ±0.2 ppm. The assigned formulas were validated by setting sensible chemical constraints (N rule; O/C

ratio ≤ 1; H/C ratio ≤ 2n+2 (C_nH_{2n+2}), double bond equivalents) in conjunction with isotope pattern comparison. Final formulas were classified into molecular groups containing CHO (blue), CHNO (orange), CHOS (green), or CHNOS (red).

Hierarchical cluster analysis based on Pearson's correlations coefficient and average linkage of assigned FT-ICR-MS peaks for both ionization modes were performed with Hierarchical Clustering Explorer Version 3.0 and R (Version 3.0.0). The assigned peaks were almost normally distributed; therefore, the usage of Pearson's correlation was adequate. This was further confirmed by hierarchical cluster analysis based on Spearman's correlation, which led to analogous clustering results. To evaluate the molecular formulas explaining the main characteristics of each subcluster, formulas with a Pearson's correlations coefficient $r > 0.9$ were extracted for the individual subclusters. These molecular formulas are characteristic for the DOM chemistry in the specific depth region of the individual subclusters. The results were visualized by the use of van Krevelen diagrams in which the hydrogen to carbon ratio (H/C) was plotted against the oxygen to carbon ratio (O/C).[4] The different bubble area represent the mean intensity of the characteristic molecular formula within the respective subcluster.

FT-ICR-MS is a semiquantitative method owing to the lack of calibration standards in nontargeted analysis and differential ionization efficiency in complex and chemically diverse mixtures. The signal intensities of mass peaks are not necessarily directly proportional to their respective concentration; however increase and decrease of mass peak intensities follow the changes in relative concentrations. To address the limited detection of some compounds due to low ionization efficiency and suppression, we performed two orthogonal ionization modes. Additionally, ionization and mass independent quantitative ^{1}H NMR and optical EEM fluorescence spectroscopy was utilized.

EEM Fluorescence Spectroscopy and PARAFAC Modeling

Excitation emission matrix fluorescence spectroscopy of SPE-DOM was measured using a Jobin Horiba Instruments (Kyoto, Japan) Aqualog spectrofluorometer. After evaporation of the methanol from 0.1 mL SPE-DOM, the samples were dissolved in NanoPure water whereby the concentration was adjusted to a maximum UV absorbance at 300 nm of 0.2. Excitation scans were conducted from 600 nm to 231 nm in 3 nm steps, and the emission wavelength ranged from 211.1 nm to 617.7 nm in ~3 nm increments. The integration time was set to 1 s. Inner filter effects and Rayleigh scattering were corrected by the Aqualog software. The spectra were normalized to a Starna® Scientific Ltd (Hainault, UK) Quinine Sulfate Standard of 1 ppm. Fluorescence units were then given in quinine sulfate units (QSU).

PARAFAC modeling was performed by the drEEM MATLAB toolbox.[5] A total of 45 EEM spectra were visually investigated, and outliers were removed. To reduce concentration effects during modeling, the samples were normalized to unit variance. After model completion, the normalization was reversed. The model evaluation and final assignment were performed as described elsewhere by taking into account the following criteria: 1) residual analysis, 2) spectral loading, and 3) split-half validation (Figure S1 and S2).[5] Fmax values describing the fluorescence intensity at the maximum for each component were exported. Principal component analysis (PCA) of Fmax values was carried out with SIMCA-P 9.0 to evaluate the influence of the components on the individual samples.

PARAFAC Model Report

Info

Toolbox	drEEM 0.1.0
Date	16.07.2015 13:54

Preprocessing

nSample - full dataset	45
nSample - modeled dataset	42
No. excluded samples	3
Excluded samples -indices	23
Scatter Removal	24
Zapped (Samples,EmRange,ExRange)	33
Fluorescence unit	35
Scaling	Normalised to unit variance in sample mode

PARAFAC model

No. PARAFAC components	3
No. Ex wavelengths	122
No. Em wavelengths	112

Validation

Split_Style	random then combine					
Split_NumBeforeCombine	3					
Split_NumAfterCombine	3					
Split_Combinations	1 2	1 3	2 3			
Split_nSample	28	28	28			
Split_AnalRuns	1	1	1			
Split_PARAFAC_options	0.000001	0	0	0	0	0
Split_PARAFAC_constraints	2	2	2			
Split_PARAFAC_convgcrit	0.000001	0.000001	0.000001			
Split_PARAFAC_Initialise	SVD					
Val_ModelName	Model3					
Val_Source	Model3it_5					
Val_Err	10679.84415					
Val_It	48					
Val_Result	Overall Result= Validated for all comparisons					
Val_Splits	AB	AC	BC			
Val_Comparisons	AB vs AC,	AB vs BC,	AC vs BC,			
Val_ConvgCrit	0.00000001					
Val_Constraints	nonnegativity					
Val_Initialise	random					

Core consist

Val_Core	89.69881479		
Val_PercentExpl	98.32132604		
Val_CompSize	63.54726548	39.62173401	28.4657282
Val_Preprocess	Reversed normalisation to recover true scores		

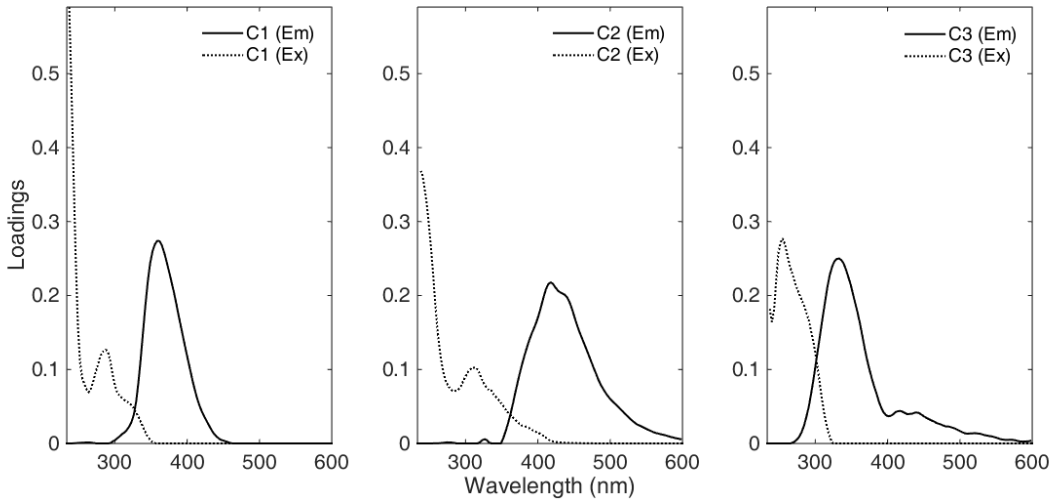

Figure S1. Excitation and emission loadings of the three components modeled with PARAFAC.

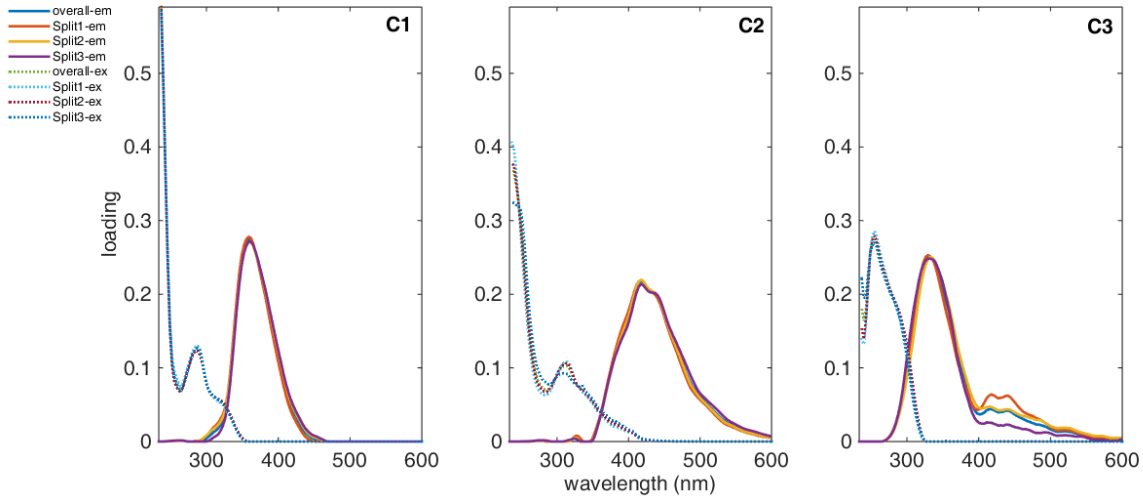

Figure S2. Split-half validation results for the three component PARAFAC model.

NMR spectroscopy

1D and 2D ^{1}H and ^{13}C NMR spectra were acquired with a Bruker (Bremen, Germany) AV III 800 spectrometer operating at B_0 = 18.7 T with Bruker standard pulse sequences and cryogenic detection. Five hundred microliters of methanolic DOM extract were dried and twice exchanged with 750 µL CD_3OD

(Merck. 99.95% ^{2}H) each before dissolution in ~130 mg CD_3OD. The solution was transferred into eventually sealed 2.5 mm Bruker MATCH™ tubes.

1D ^{1}H NMR spectra were recorded with a spin-echo sequence (10 µs delay) to allow for high-Q probe ringdown and classical presaturation to attenuate residual water present "*noesypr1d*" (typically 4–16 k scans with 5 s acquisition time, 5 s relaxation delay, 1 ms mixing time; 1 Hz exponential line broadening). Integration of the spectra was performed in the chemical shift range δ_H = 0.5–9.5 ppm with the exclusion of residual water (δ_H = 4.9–5.1 ppm) and methanol (δ_H = 3.25–3.35 ppm) NMR resonances by means of AMIX-based bucket analysis (0.1 ppm uniform width, normalized total ^{1}H NMR integral = 100%). The integrals were further grouped according to their specific chemical regions of aromatics (δ_H = 7.0–9.5 ppm, green), olefins (δ_H = 5.1–7.0 ppm, blue), carbohydrates (δ_H = 4.9–3.1 ppm, pink), CRAM (δ_H = 3.1–1.9 ppm, purple), and aliphatics (δ_H = 0.5–1.9 ppm, orange) (cf. Figure 5, Table S3 and also Tables S4 and S5, with more detailed attribution of key substructures to δ_H ranges).

The one bond coupling constant $^{1}J(CH)$ used in 2D $^{1}H,^{13}C$ DEPT-HSQC spectra (*hsqcedetgpsisp2.2*) was set to 145 Hz; other conditions were as follows: ^{13}C 90 degree decoupling pulse, GARP (70 µs); 50 kHz WURST 180 degree ^{13}C inversion pulse (Wideband, Uniform, Rate, and Smooth Truncation; 1.2 ms); F2 (^{1}H): spectral width of 5981 Hz (11.96 ppm); 1.25 s relaxation delay; F1 (^{13}C): SW = 17607 Hz (140 ppm). Absolute value JRES and echo-antiecho TOCSY spectra (with solvent suppression: *jresgpprqf*, *dipsi2etgpsi19*) used a spectral width of 5498 Hz [JRES (F1) = 50 Hz] and were computed to a 16384 × F1 matrix [JRES/TOCSY (F1) = 128/4096]. Other NMR acquisition conditions are given in Supporting Information Table S1.

Table S1. Acquisition conditions for NMR spectra provided in Figures 5 and S7-8. PK: NMR probehead used, 8Q: 800 MHz 5 mm cryogenic inverse quaternary $^{1}H/^{13}C/^{15}N/^{31}P$; NS: number of scans (for 2D NMR: F2); AQ: acquisition time [ms]; D1: relaxation delay [ms]; NE: number of F1 increments in 2D NMR spectra; WDW1, WDW2: apodization functions in F1/F2 (EM exponential line broadening factor [Hz]; SI: sine bell); PR1, PR2: coefficients used for windowing functions WDW1, WDW2, EM is given in [Hz], SI derived functions indicate shift by π/n. Total NMR acquisition time AQ_{Σ} is computed as follows: $AQ_{\Sigma} = NS \times (D1 + AQ) \times NE$, with NE = 1 for 1D NMR spectra.

Experiment	**Figure**	**PK**	**NS**	**AQ [ms]**	**D1 [ms]**	**NE**	**WDW1**	**WDW2**	**PR1**	**PR2**
^{1}H	5, S6	8Q	480--15616	5000	5000	/	/	EM	-	1
JRES	S7	8Q	2048	1000	500	64	QS	QS	0	0
TOCSY	S7, S8	8Q	160	1000	500	2243	QS	EM	6	2.5
DEPT-HSQC	S8	8Q	2048	250	1250	317	QS	EM	2.5	2.5

Results and Discussion

FT-ICR-MS of DOM

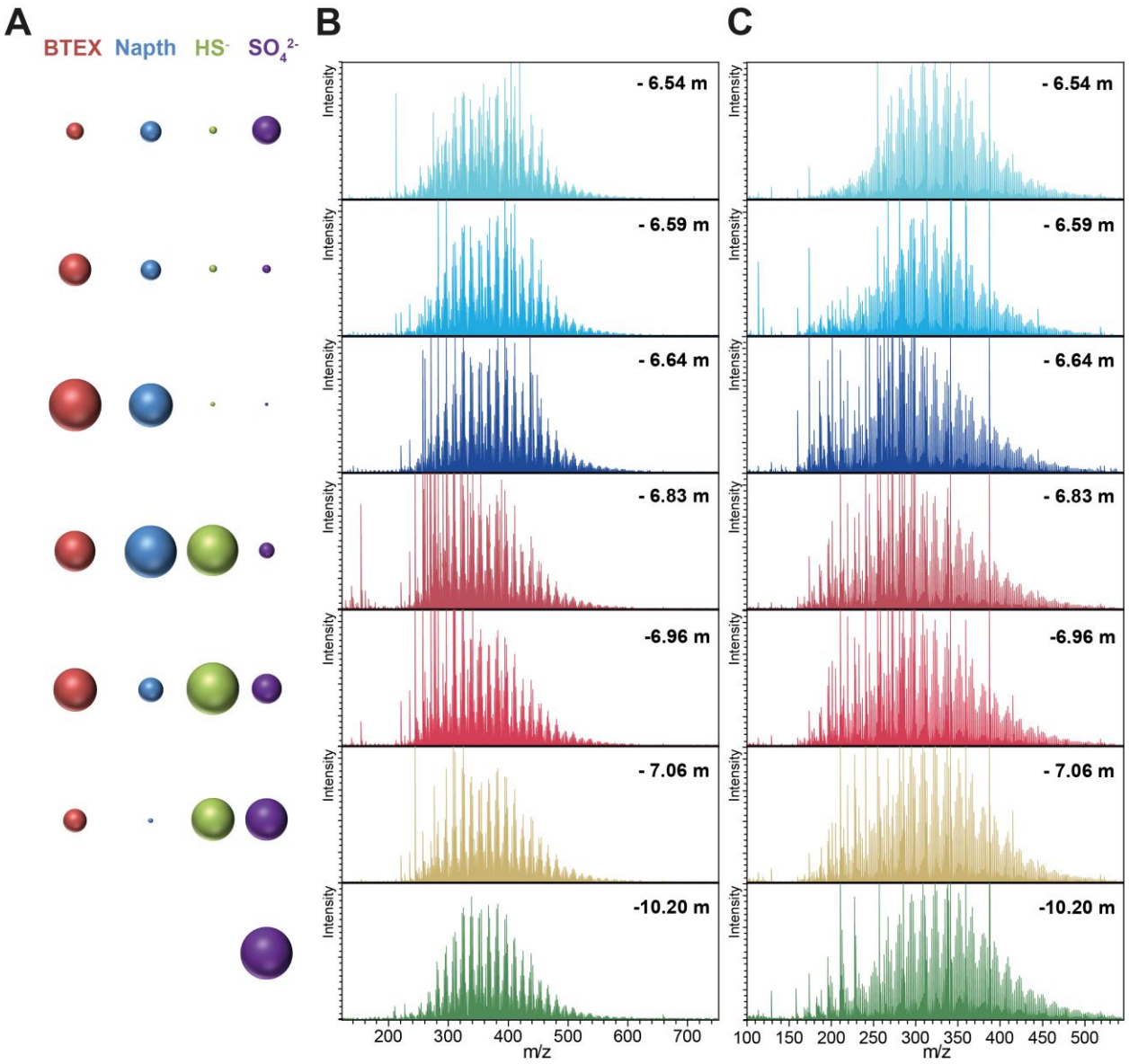

Figure S3. (A) The bubbles display the BTEX (red), naphthalene (blue), hydrogen sulfide (light green), and sulfate (purple) concentrations in the groundwater. (B) (-)ESI FT-ICR mass spectra from m/z 123–750 and (C) (+)APPI FT-ICR mass spectra from m/z 100–550 of selected SPE-DOM samples along the aquifer.

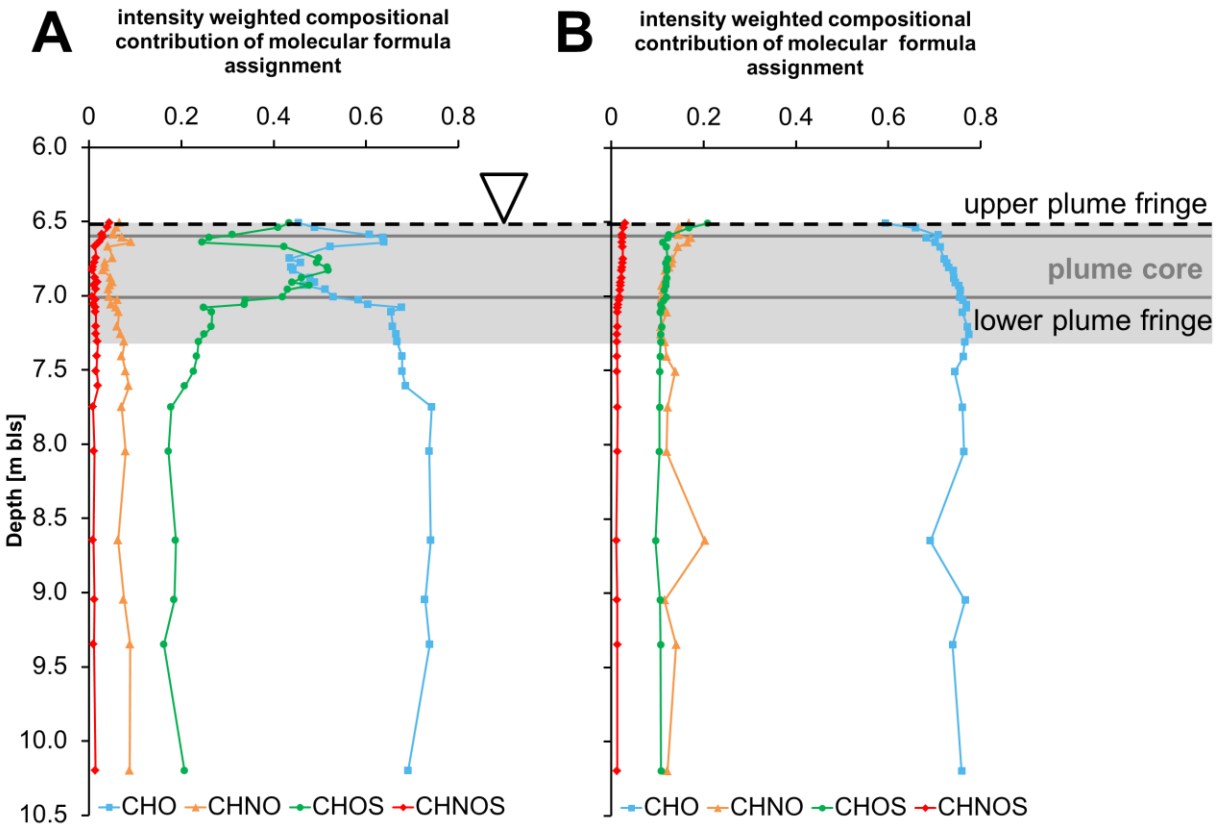

Figure S4. (A) (-)ESI and (B) (+)APPI FT-ICR mass spectra derived intensity weighted compositions of assigned molecular formulas showed increased contribution of CHOS and decreased CHO compounds in the highly sulfidic region from 6.67–7.06 m bls in (-)ESI FT-ICR mass spectra. (+) APPI FT-ICR mass spectra did not show this trend, thereby suggesting that CHOS compounds were formed from highly functionalized CHO molecules as a result of bacterial sulfate reduction.

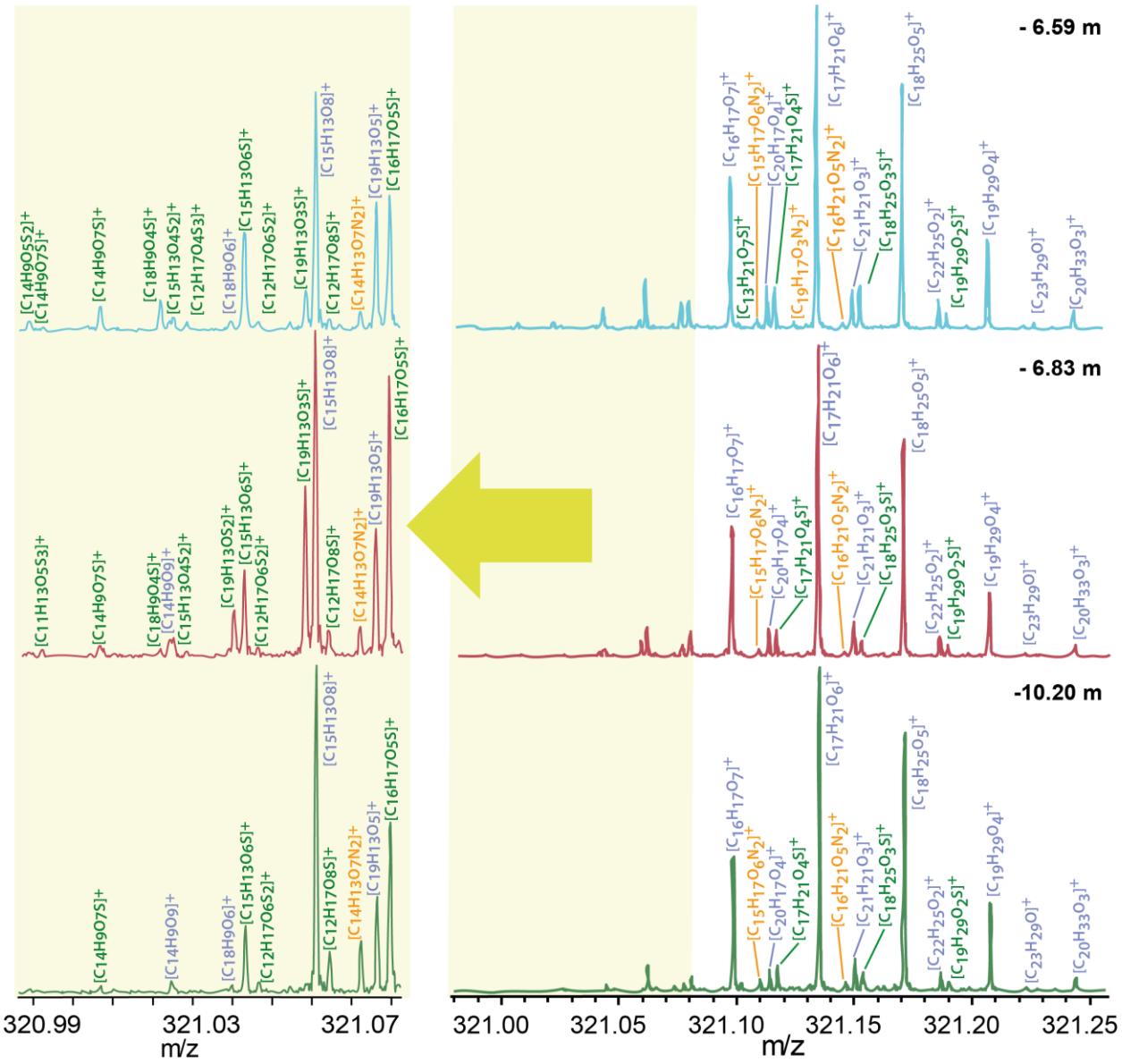

Figure S5. Expansion of nominal mass 321 from (+)APPI FT-ICR mass spectra of selected SPE-DOM samples and enhanced relative intensity for the mass segment m/z 320.9–321.08, which indicates that CHOS molecules are discriminating the samples where bacterial sulfate reduction is highly active (6.59 m bls and 6.83 m bls) from the bottom sample (10.20 m bls).

EEM fluorscence spectroscopy of DOM

Table S2. EEM fluorescence spectroscopy in combination with PARAFAC modeling results compared to other studies.

component	Ex λ [nm]	Em λ [nm]	description	Zhou et al. 2012[6]	Mendoza et al. 2013[7]	Zhou et al. 2013[8]	Bianchi et al. 2014[9]
				Comp: Ex/Em (nm) (description)			
C1	237 (288)	358.7	oil, BTEX, PAH and its degradation products	C3: 232/346 (oil)	C4 & C6: 225, 270, 280/340 (naphthalene-like enriched, benzene/arene-like)	C2: 236/350 (oil - related, degradation product)	
C2	255 (315)	417.6	humic-like	C4: 248/446 (humic like)	C3 250/440 (humic-like)	C3: 256 (340)/460 (terrestrial humic substance, and chemically dispersed oil)	C1: 240/400-436 (terrestrial humic substance-like)
C3	255	332.8	oil, BTEX, PAH and its degradation products	C2: 264/234 (oil)	C2: 220, 255, 270/330, (benzene/arene-like enriched, naphthalene-like)		

NMR spectroscopy of DOM

Table S3. ^{1}H NMR section integrals derived from total area-normalized ^{1}H NMR spectra integrals with main structures provided for selected samples along the aquifer, according to sampling depth.

structure	key structures	δ(^{1}H) [ppm]	6.54 [m bls]	6.59 [m bls]	6.64 [m bls]	6.83 [m bls]	6.96 [m bls]	7.06 [m bls]	10.2 [m bls]
aromatics	$\underline{\mathbf{H}}_{ar}$	9.5 - 7.0	14.1	20.1	25.9	29.9	25.5	19.3	12.1
olefins	$\underline{\mathbf{H}}$C=C, $\mathbf{H}CO_2$	7.0 - 5.1	3.6	4.7	3.8	3.4	2.5	2.6	2.3
carbohydrates	$\underline{\mathbf{H}}$CO	4.9 - 3.1	13.0	11.8	10.0	10.7	10.4	12.3	12.4
CRAM	$\underline{\mathbf{H}}$CCX	3.1 - 1.9	32.7	32.0	32.5	29.9	31.1	31.2	31.9
aliphatics	$\underline{\mathbf{H}}$CCC	1.9 - 0.5	36.5	31.5	27.8	26.0	30.4	34.7	41.2

Table S4. ^{1}H NMR section integrals derived from total area-normalized ^{1}H NMR spectra integrals with main structures provided for selected samples along the aquifer, according to sampling depth (cf. Table S5 for attribution of δ_H section integrals).

structure		$\delta(^{1}H)$ [ppm]	6.54 [m bls]	6.59 [m bls]	6.64 [m bls]	6.83 [m bls]	6.96 [m bls]	7.06 [m bls]	10.2 [m bls]
aromatics **H**$_{ar}$	a	9.5 - 8.9	0.2	0.3	0.2	0.1	0.2	0.2	0.1
	b	8.9 - 8.3	1.0	1.2	1.6	1.4	1.4	1.1	0.9
	c	8.3 - 7.3	8.1	11.7	16.4	19.1	15.4	11.7	7.5
	d	7.3 - 7.0	4.7	6.9	7.7	9.3	8.5	6.4	3.5
olefins **H**C=C, **H**CO_2	e	7.0 - 6.5	1.8	2.8	3.2	2.8	2.5	2.1	1.7
	f	6.5 - 6.0	0.5	0.8	0.4	0.3	0.0	0.1	0.4
	g	6.0 - 5.1	1.4	1.1	0.2	0.4	0.0	0.5	0.2
carbohydrates **H**CO	h	4.9 - 3.1	13.0	11.8	10.0	10.7	10.4	12.3	12.4
CRAM **H**CCX	i	3.1 - 2.1	27.3	27.4	28.6	26.4	27.2	26.8	26.7
	j	2.1 - 1.9	5.4	4.6	3.9	3.6	3.9	4.4	5.2
aliphatics **H**CCC	k	1.9 - 1.3	20.4	16.8	14.6	13.7	15.4	18.3	22.4
	l	1.3 - 1.2	4.5	4.0	4.0	3.6	4.7	5.4	6.4
	m	1.2 - 0.5	11.6	10.6	9.1	8.7	10.3	11.0	12.4

243 Table S5. Description of ^{1}H NMR key structural units of DOM and PAH degradation products observed
244 in this study.

	$\delta(^{1}H)$ [ppm]	**key structural units of DOM and PAH degradation products observed in this study**
a*	9.5–8.9	multi (≥4) ring condensed aromatics
b*	8.9–8.3	multi (≥3) ring condensed aromatics
c*	8.3–7.3	single and two-ring aromatics; six-membered N-heterocycles; oxidized aromatics, i.e., mainly (poly)carboxylated aromatics
d*	7.3–7.0	alkylated and non-substituted monoaromatic rings
e*	7.0–6.5	aromatics with oxygenated substituents (OH, OR); alkylated aromatics with fused alicyclic units attached
f	6.5–6.0	conjugated double bonds: =C-C=C**H**; five membered ring heterocycles (O, N, S)
g	6.0–5.1	isolated double bonds: =C**H**; anomerics in carbohydrates: O_2C**H**
h	4.9–3.1	OC**H** oxygenated aliphatics (e.g., carbohydrates, esters, ethers)
i*	3.1–2.1	remotely functionalized aliphatics: OCC_{β}**H**, HOOC-C_{α}**H**-C (aliphatic carboxylic acids); δ_H > 2.5 ppm: C_{ar}-C**H**-C**H**-COOH, -C-C**H**-NH
j	2.1–1.9	acetate, remotely functionalized aliphatics: OCCC**H**
k	1.9–1.3	OCCC**H**, branched aliphatics, fused alicyclic rings
l	1.3–1.2	$(C\mathbf{H}_2)_n$ polymethylene; certain branched aliphatics, alicyclic rings
m	1.2–0.5	certain branched aliphatics: CCCC**H**, $C\mathbf{H}_3$ groups, certain alicyclic rings

* occurrence of PAH and its degradation products; other NMR resonances mainly comprise SPE-DOM structures

245

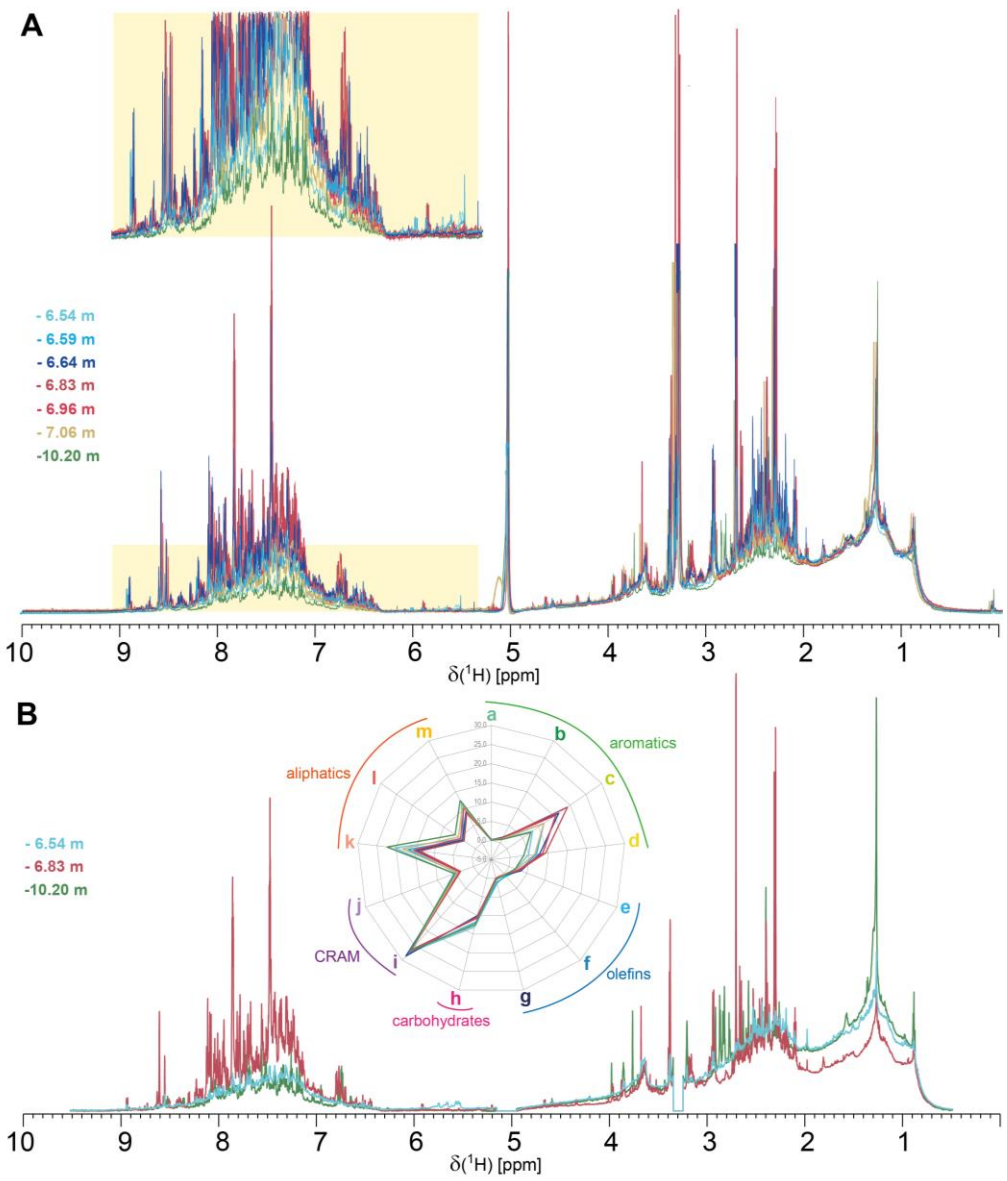

Figure S6. Overlay of ^{1}H NMR spectra of SPE-DOM (800 MHz, CD_3OD). (A) Visual overlay according to the common intensity of "pure" aliphatics (CCC<u>H</u>; δ_H = 0.5–1.9 ppm) revealed the limited variance of the DOM bulk signature even in the case of highly petroleum contaminated SPE-DOM. (B) Overlay of selected ^{1}H NMR spectra, which were normalized to the total integral area, that show the increased integral area in the aromatic region for contaminated SPE-DOM (6.83 m bls) compared to the increased integral of the aliphatic region for lesser petroleum contaminated SPE-DOM from the very top (6.54 m bls) and bottom (10.20 m bls) of the aquifer. The radar diagram of the ^{1}H NMR section integrals, derived from total area-normalized ^{1}H NMR spectra integrals according to structural zones, pointed out the increased aromatic proportion for contaminated SPE-DOM compared to increased aliphatic integral proportion for lesser petroleum contaminated SPE-DOM.

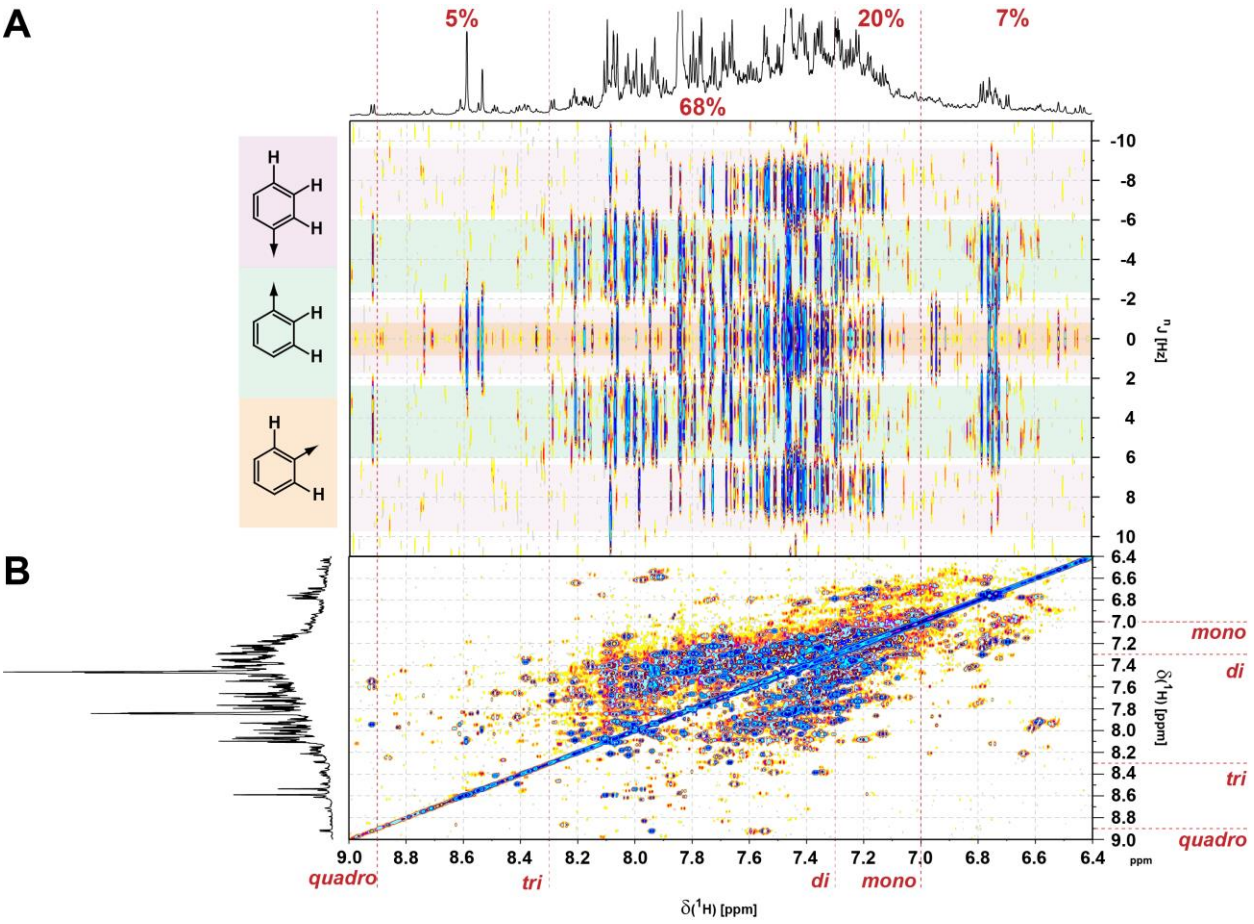

Figure S7. (A) JRES, and (B) TOCSY NMR spectra (800 MHz, CD_3OD) of SPE-DOM taken at 6.83 m bls. Joint positioning of δ_H and $^nJ_{HH}$ indicated the absence of olefinic protons in accordance with negligible ^1H NMR integral at δ_H ~ 5.15–6.3 ppm (Figure 5, S6 and Table S4). An increased ring count in alkylated polycyclic hydrocarbons induced progressive downfield chemical shift, while cumulative carboxylation behaved analogously. The projection NMR spectrum is provided with aromatic NMR section integral (sum = 100%) for various fused ring arrangements (cf. Table S3). Panel A: Triplett splittings with $^3J_{HH} \approx 8$ Hz indicated HC_{ar}-**H**C_{ar}-**H**C_{ar} groups (purple shade; 1,2,3 and 1,2,3,4-substitution); dublett splittings with $^3J_{HH} \approx 8$ Hz indicated isolated ortho C_q-**H**C_{ar}-**H**C_{ar}-C_q groups (green shade), and complex splittings with $^4J_{HH} < 3$ Hz indicated isolated meta protonated **H**C_{ar}-C_q-**H**C_{ar} groups (orange shade). Panel B: TOCSY cross peaks from vicinal intra aromatic correlations (**H**C_{ar}-**H**C_{ar} : $^3J_{HH} \approx$ 8 Hz) with chemical shift ranges for singly ($\delta_H > 7.0$ ppm), doubly ($\delta_H > 7.3$ ppm), and higher fused aromatic systems ($\delta_H > 8.3$ ppm) indicated in color. Alicyclic rings condensed with aromatic rings caused the latter to resonate below 7 ppm ($\delta_H < 7$ ppm); likewise, oxygenated aromatics did the same. However, the latter are at best very marginal contributors and will not account for the appreciable NMR cross peak integral observed here. The observed ^1H NMR chemical shift and cross peak distribution suggested the presence of mostly substituted single and doubly fused aromatic rings.

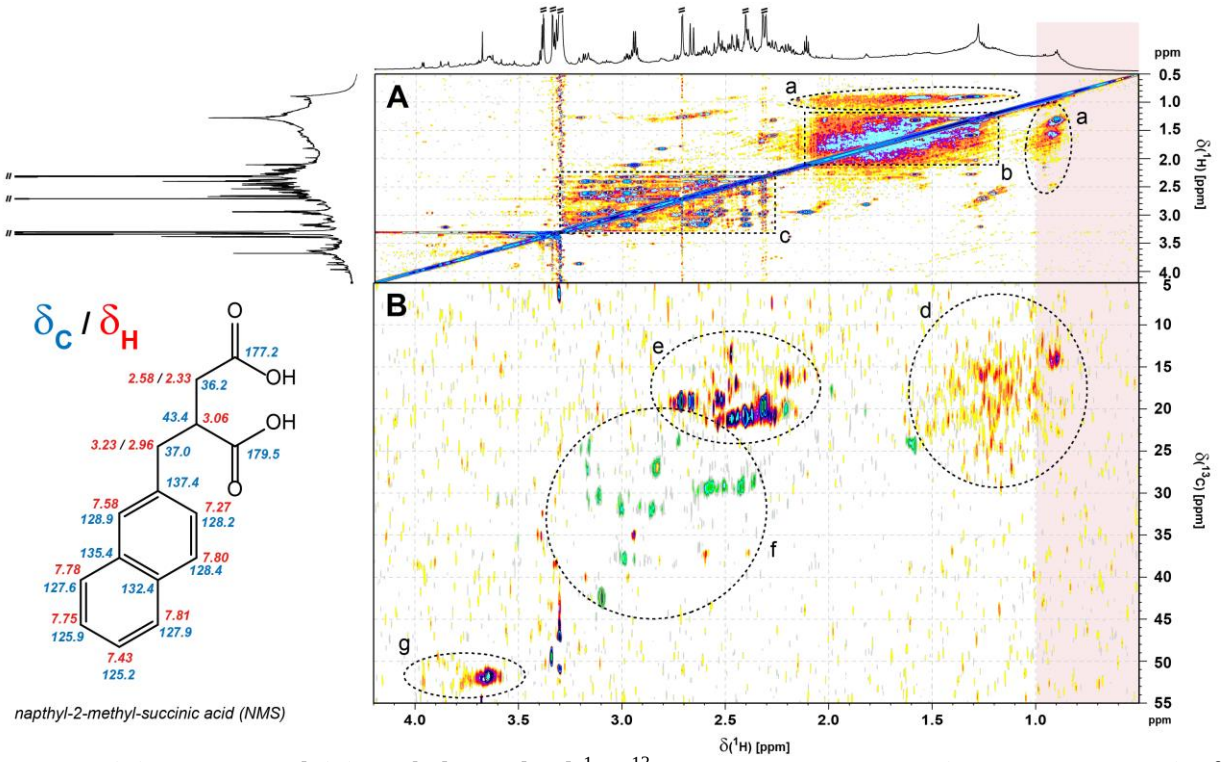

Figure S8. (A) TOCSY and (B) methylene-edited ^{1}H, ^{13}C HSQC NMR spectra (800 MHz, CD_3OD) of aliphatic ^{1}H chemical shift region (δ_H = 0.5–4.2 ppm) from SPE-DOM sample taken at 6.83 m bls. Panel A: section a: C-CH-C**H**$_3$ cross peaks; section b: intra-aliphatic C-**CH**-C$_n$-**CH**-C cross peaks (n = 0, 1), section c: C$_{ar}$-**CH**-C$_n$-**CH**-COOH cross peaks (n = 0, 1). Panel B: (B) general colors of cross peaks: CH_3, CH: red, and CH_2: green. Section d: C-**CH$_3$** cross peaks; section e: C$_{ar}$-**CH$_3$** cross peaks; section f: C$_{ar}$-**CH$_2$**-C and C$_{al}$-**CH$_2$**-COOH cross peaks; section g: methyl esters **H$_3$C**O-C(=O)-C-. Cross peaks at δ_H = 2.0–3.25 ppm were indicative for protons in α-positons to carboxylic groups (**H**Cα-COOH) and those attached to aromatic groups **H**Cα-C$_{ar}$. Both TOCSY and ^{1}H, ^{13}C DEPT HSQC spectra of aliphatic spin systems indicated the presence of structural subunits as found in common PAH degradation products such as succinic acid derivatives like NMS (shown in the insert with $\delta_{H/C}$ as computed from ACD Labs software).[10-12]

References

1. Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G., A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr.: Methods* **2008,** *6*, 230-235; DOI 10.4319/lom.2008.6.230.
2. Podgorski, D. C.; McKenna, A. M.; Rodgers, R. P.; Marshall, A. G.; Cooper, W. T., Selective ionization of dissolved organic nitrogen by positive ion atmospheric pressure photoionization coupled with Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2012,** *84* (11), 5085-5090; DOI 10.1021/ac300800w.
3. Tziotis, D.; Hertkorn, N.; Schmitt-Kopplin, P., Letter: Kendrick-analogous network visualisation of ion cyclotron resonance Fourier transform mass spectra: improved options for the assignment of elemental compositions and the classification of organic molecular complexity. *Eur. J. Mass.* **2011,** *17* (4), 415-421; DOI 10.1255/ejms.1135.
4. van Krevelen, D. W., Graphical-statistical method for the study of structure and reaction processes of coal. *Fuel* **1950,** *29*, 269-284.
5. Murphy, K. R.; Stedmon, C. A.; Graeber, D.; Bro, R., Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Anal. Methods* **2013,** *5* (23), 6557–6882; DOI 10.1039/c3ay41160e.
6. Zhou, Z.; Guo, L., Evolution of the optical properties of seawater influenced by the Deepwater Horizon oil spill in the Gulf of Mexico. *Environ. Res. Lett.* **2012,** *7* (2), 1-12; DOI 10.1088/1748-9326/7/2/025301.
7. Mendoza, W. G.; Riemer, D. D.; Zika, R. G., Application of fluorescence and PARAFAC to assess vertical distribution of subsurface hydrocarbons and dispersant during the Deepwater Horizon oil spill *Environ. Sci.: Processes Impacts* **2013,** *15* (5), 1017-1030; DOI 10.1039/c3em30816b.
8. Zhou, Z.; Guo, L.; Shiller, A. M.; Lohrenz, S. E.; Asper, V. L.; Osburn, C. L., Characterization of oil components from the Deepwater Horizon oil spill in the Gulf of Mexico using fluorescence EEM and PARAFAC techniques. *Mar. Chem.* **2013,** *148*, 10-21; DOI 10.1016/j.marchem.2012.10.003.
9. Bianchi, T. S.; Osburn, C.; Shields, M. R.; Yvon-Lewis, S.; Young, J.; Guo, L.; Zhou, Z., Deepwater horizon oil in gulf of Mexico waters after 2 years: transformation into the dissolved organic matter pool. *Environ. Sci. Technol.* **2014,** *48* (16), 9288-9297; DOI 10.1021/es501547b.
10. Jobelius, C.; Ruth, B.; Griebler, C.; Meckenstock, R. U.; Hollender, J.; Reineke, A.; Frimmel, F. H.; Zwiener, C., Metabolites indicate hot spots of biodegradation and biogeochemical gradients in a high-resolution monitoring well. *Environ. Sci. Technol.* **2011,** *45* (2), 474-481; DOI 10.1021/es1030867.
11. Griebler, C.; Safinowski, M.; Vieth, A.; Richow, H.; Meckenstock, R. U., Combined Application of Stable Carbon Isotope Analysis and Specific Metabolites Determination for Assessing In Situ Degradation of Aromatic Hydrocarbons in a Tar Oil-Contaminated Aquifer. *Environ. Sci. Technol.* **2004,** *38* (2), 617–631; DOI 10.1021/es0344516.
12. Jarling, R.; Kuhner, S.; Basilio Janke, E.; Gruner, A.; Drozdowska, M.; Golding, B. T.; Rabus, R.; Wilkes, H., Versatile transformations of hydrocarbons in anaerobic bacteria: substrate ranges and regio- and stereo-chemistry of activation reactions. *Front. Microbiol.* **2015,** *6* (880), 1-14; DOI 10.3389/fmicb.2015.00880.