

# **High field NMR Spectroscopy and FTICR Mass Spectrometry: Powerful Discovery Tools for the Molecular Level Characterization of Marine Dissolved Organic Matter**

N. Hertkorn<sup>1\*</sup>, M. Harir<sup>1</sup>, B. P. Koch<sup>2</sup>, B. Michalke<sup>1</sup>, Ph. Schmitt-Kopplin<sup>1</sup>

<sup>1</sup> Research Unit Analytical Biogeochemistry, Helmholtz Zentrum München, Ingolstädter Landstrasse 1, 85758 Neuherberg, Germany

\*Corresponding author. Tel: +4989-3187-2834; fax: +4989-3187-2705.

E-mail address: hertkorn@helmholtz-muenchen.de (Norbert Hertkorn);

<sup>2</sup> Alfred Wegener Institute, AWI, Am Handelshafen 12, D-27570 Bremerhaven, (Building Co-5), Germany

## **SUPPLEMENT**

### **SI 1 transfer functions referring to organic structural spectroscopy of marine SPE-DOM (NMR and FTMS)**

The two most significant methods of non-target organic structural spectroscopy of marine SPE-DOM also show the most clearly defined relationships between intrinsic molecular properties and measured variable (Hertkorn et al., 2007). NMR spectroscopy detects the precession frequencies of individual magnetic moments of dissolved SPE-DOM atoms in an external magnetic field  $B_0$ . Quantitative relationships between number of spins and area of NMR resonances operate in absence of differential NMR relaxation.

The spacings between NMR energy levels (Zeeman splittings) are minute and hence, NMR relaxation is slow compared with that of other molecular energy levels. The rate of NMR relaxation depends on the chemical environment of the respective nuclei and nature of their motion. Slow longitudinal NMR relaxation (cf. Table S1) defines a lower limit on the rate at which an NMR experiment can be repeated: equilibrium magnetization has to be re-established in-between successive scans; otherwise, respective NMR resonances will be reduced in area (Keeler, 2005). Very fast transverse NMR relaxation in marine SPE-DOM molecular segments with large chemical shift anisotropy and/or low internal mobility might produce excessive line widths, and eventually NMR resonances too broad for observation.

Quantification uncertainties in FTMS spectra of marine DOM are potentially large and primarily refer to ionization selectivity of individual marine SPE-DOM molecules and massive intrinsic averaging (projection of many isomers onto single mass peaks; see also section 3.12).

**Table S1.** Transfer functions of NMR and FTICR mass spectrometry interfering with accurate quantification of marine SPE-DOM

method	atomic, molecular, physical and chemical process	susceptible substructures in marine SPE-DOM
NMR	quantitative relationships between number of spins and area of NMR resonances operate in absence of differential NMR relaxation	all marine SPE-DOM substructures (cf. below)
NMR	low intrinsic overall NMR sensitivity restricts accessible S/N ratio and accuracy of signal definition	all substructures; depends on NMR experiment employed and refers primarily to hydrogen deficient units in heteronuclear NMR spectroscopy (cf. Table S2)
NMR	longitudinal (spin-lattice) relaxation of z-magnetization (energy transfer with longitudinal relaxation time $T_1$ )	expedient longitudinal relaxation applies for hydrogen-rich chemical environments; slow longitudinal relaxation applies for hydrogen-deficient molecular environments in marine SPE-DOM (alkylated fused rings, condensed aromatics, largely unsaturated molecules, possibly with high content of heteroatoms: O, N, P, S); scales roughly with gyromagnetic ratio, $T_1$ : $^1\text{H} < ^{13}\text{C} \sim ^{31}\text{P} < ^{15}\text{N}$ ; very efficient paramagnetic relaxation mediated by the large electron magnetic moment ( $\mu_e \sim 658.2 \mu_H$ ) occurs in organic radicals and paramagnetic metal coordination compounds (e.g. with 3d-transition metals, like $\text{Fe}^{2+}$ and $\text{Fe}^{3+}$ ) and might eventually broaden NMR resonances of proximate atoms beyond recognition ( $T_1 \approx T_2$ ) (Bertini et al., 2008)
NMR	transverse (spin-spin) relaxation of xy-magnetization (loss of phase coherence in observation plane) governs NMR line width: $\Delta\nu \sim 1/\pi \cdot T_2$	fast transverse NMR relaxation primarily operates in highly anisotropic chemical environments (e.g. $\text{sp}^2$ -hybridized carbon in marine SPE-DOM). While it might not regularly broaden NMR resonances in 1D NMR spectra of marine SPE-DOM beyond recognition, it will attenuate respective 2D NMR cross peaks at higher F1-resolution (cf. HSQC NMR spectra acquired at different F1 resolution: Fig. 9 and Fig. 13)
FTICR /MS	ionization in mass spectrometry is a very complex and ill-defined succession of distinct but interdependent events which effect the transformation of dissolved marine SPE-DOM molecules into gas-phase marine SPE-DOM ions, which are detected in high vacuum; any compositional and structural	ionization selectivity in mixtures follows mass (decreasing volatility and likely surface activity impedes susceptibility for ionization from liquid phase droplets) and functional group content of respective molecules. In complex mixtures, ionization selectivity decisively depends on the ionization method applied; lesser statistical probability to observe individual compositions $\text{C}_n\text{H}_m\text{O}_q\text{Z}_\Sigma$ ( $\text{Z}_\Sigma$ : all combinations of heteroatoms) within given nominal mass ranges is expected at higher mass. More numerous mathematical combinations to form different

	selectivity within this entire sequence will interfere with quantification of complex mixtures in mass spectra	molecular compositions will become available at higher mass, and the mandatory dissimilarity with respect to compositions and structures within clusters of nominal mass will tremendously increase (Hertkorn et al., 2008). The count of feasible isomers for any given composition will increase even more massively with increasing nominal mass (Hertkorn et al., 2007). However, large marine SPE-DOM molecules appear to be more closely related to biomolecules than small ones (Amon and Benner, 1996), which will be less aged and processed and which then will show limited compositional variance in comparison with more evolved biogeochemical mixtures
FTICR /MS	mass spectra project the entire structural space on a highly resolved mass axis. The resulting compositional space represents the isomer-filtered complement of the entire space of molecular structures and is quantized according to the laws of chemical binding (Hertkorn et al., 2008); the conversion of the entire structural space into a mass spectrum amounts to an extreme data reduction, while crucial molecular-level information (i.e. the knowledge of the molecular composition) is retained.	any mass peak within FTICR mass spectra of marine SPE-DOM will represent very many isomers; this extreme data reduction applies throughout the entire mass spectrum. The count of chemically reasonable isomers for any given mass is maximal at average H/C and O/C elemental ratios (Hertkorn et al., 2007 and 2008). In general, mass peak intensities in FTICR mass spectra of marine SPE-DOM follow the counts of feasible isomers for given compositions; for strongly hydrogen (and oxygen) deficient molecular compositions (as well as for near saturated ones), reasonable proposals of principal structures can be deduced from molecular compositions because the accessible structural diversity is severely restricted by common laws of chemical binding (Koch and Dittmar, 2006)
FTICR /MS	maximum coverage of the theoretical C,H,O-compositional space at average H/C and O/C ratios is readily observed in high quality FTICR mass spectra. Hence, this “ultimate mass resolution” in mass spectra of marine SPE-DOM compromises the discrimination capacity of mass spectra. Here, emphasis has to be placed on more reliable overall quantification and definition of heteroatom-containing molecules $C_nH_mO_qZ_\Sigma$ ( $Z_\Sigma$ : all combinations of heteroatoms)	massive intrinsic averaging preferentially refers to elemental compositions at average H/C and O/C elemental ratios (Hertkorn et al., 2007)

## SI 2 Intrinsic aspects of high-field NMR spectroscopy related to sensitivity and resolution in complex mixtures

Solution NMR spectroscopy at 800 MHz ( $B_0 = 18.8$  T) with cryogenic detection offers ultimate sensitivity and dispersion (Kovacs et al., 2005). Both items are very useful for analysis of marine NOM, an extremely intricate organic mixture, which occurs at low concentrations ( $< 1$  mg/L seawater) and shows low-resolution signatures in one-dimensional proton and carbon NMR spectra because of prevalent superposition of many NMR resonances at any data point (Hertkorn et al., 2007). Notably, individual NMR experiments carry specific intrinsic nominal resolution. Here, two-dimensional NMR spectra not only exceed the nominal resolution of 1D NMR spectra considerably (Table S2), but 2D NMR pulse sequences also act as filters to selectively emphasize individual transfers of magnetization such as certain homonuclear and heteronuclear spin-spin couplings (Cavanagh et al., 2007; Hertkorn et al., 2002b). Hence, 2D NMR spectra of NOM commonly exhibit a vastly superior effective resolution even if they might represent a lesser overall number of NMR resonances than those which define a 1D NMR spectrum. Analogously, individual NMR experiments will carry intrinsic characteristics with respect to spectral resolution or/and S/N ratio which can be selectively pronounced and attenuated by judicious choice of acquisition and apodization parameters (Hoch and Stern, 1996).

Increased NMR detection sensitivity at high magnetic field  $B_0$  might elevate potential interferences intrinsic to NMR to the level of observation which otherwise would be lost in noise. NMR spectra represent operating quantum mechanics with macroscopic observables, which carry information about chemical environments of atoms as well as of physical phenomena, like atomic and molecular mobility (Carper et al., 2004; Case, 1998; Cook et al., 1996; Effemey et al., 2000; Ying et al., 2006). The different NMR experiments will react unequal to these challenges. For example, NMR experiments with long intrinsic duration and variable magnetization transfer / relaxation delays (like 2D NMR experiments) will be more susceptible to effects of fast and differential relaxation than rather short duration NMR pulse sequences (like 1D NMR experiments).

The Larmor equation of NMR [ $\omega_i = \gamma_N \cdot B_0 \cdot (1 - \delta_i)$ ;  $\gamma_N =$  gyromagnetic ratio] implies that actual NMR frequencies  $\omega_i$  are proportional to the external magnetic field  $B_0$ . Likewise, NMR total chemical shielding ranges as well as frequency differences  $\Delta\omega_{ij}$  between dissimilar chemical environments  $\omega_i$  and  $\omega_j$  scale with  $B_0$ . Nevertheless, NMR spectra are visualized as plots of line intensity versus frequency  $\omega_i$ , expressed as dimensionless, magnetic field independent units of chemical shift  $\delta_i$ . Hence, the effects of differential NMR relaxation which itself depend on atomic and molecular mobility on frequency timescales (Korzhnev et al., 2004a; Neudecker et al., 2009) are magnetic field dependent and proportional to  $B_0^m$  ( $m \geq 1$ , depending on the actual relaxation mechanism; Bakhmutov, 2004).

NMR relaxation strongly affects NMR spectra: fast transverse relaxation directly relates to line broadening in one and two dimensional NMR spectra (Cavanagh et al., 2007). In the latter, fast transverse relaxation will attenuate cross peak intensities at higher F1 increments, resulting in lesser overall S/N ratio at long acquisition times and often noticeable diminished F1 resolution. Two major molecular mechanisms induce fast transverse NMR relaxation: larger molecules exhibit higher proportions of low frequency atomic and molecular motions which impose efficient transverse relaxation (Hansen and Al-Hashimi, 2007; Korzhnev et al., 2004b). In case of DOM, effects of differential transverse NMR relaxation, which might lead to significant line broadening are rather expected for molecular units of low mobility and low local symmetry (e.g.  $sp^2$ -hybridized carbon; Buddrus

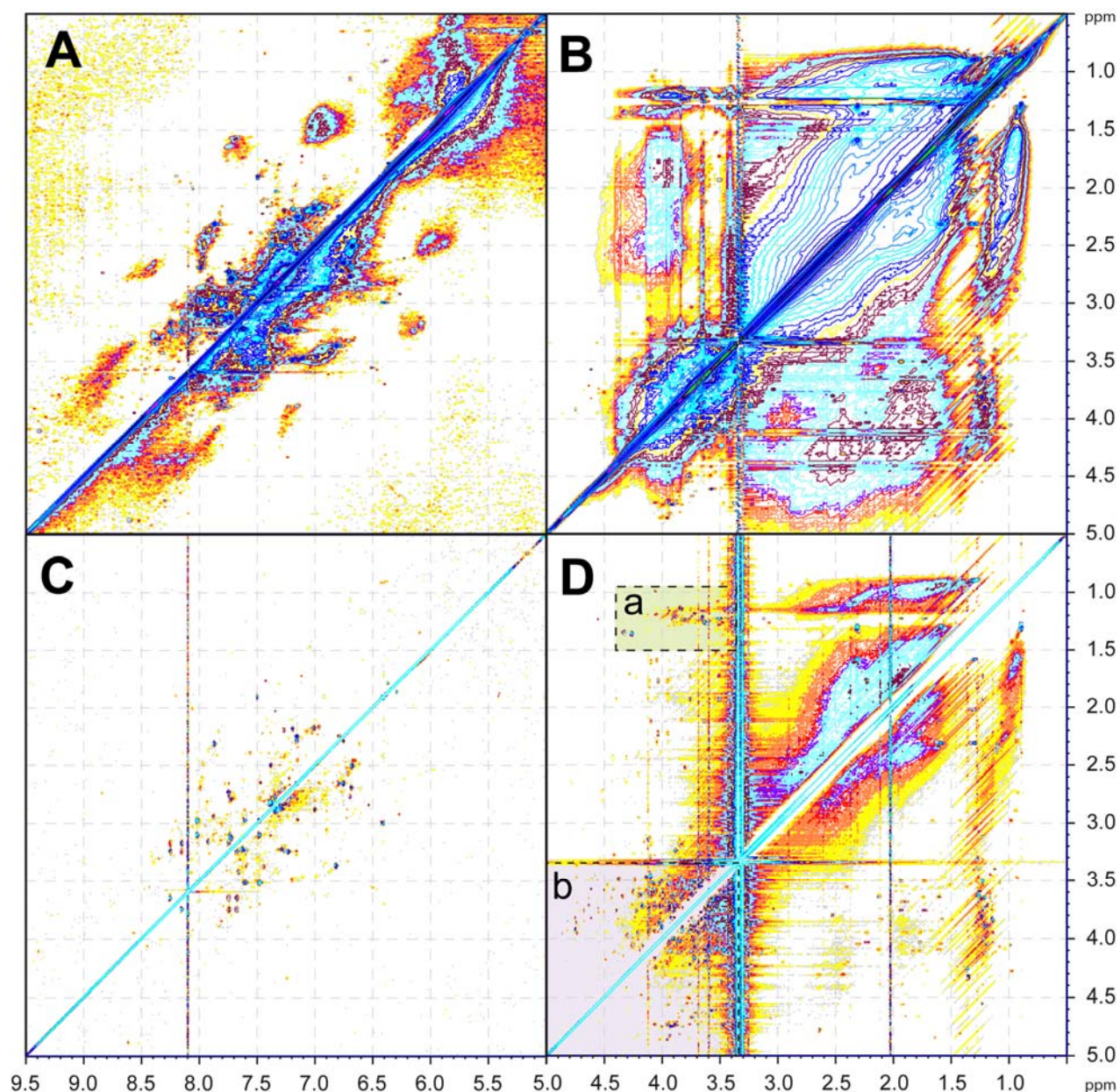
et al., 1989). All  $sp^2$ -hybridized carbon chemical environments in marine DOM will be subject to fast NMR relaxation at high magnetic field (here:  $B_0 = 18.8$  T) because of a significant contribution of chemical shift anisotropy (CSA) to transverse relaxation (Carper et al., 2004; Effemey et al., 2000; Hansen and Al-Hashimi, 2007; Korzhnev et al., 2004b; Mulder and Akke, 2003). Therefore, marine DOM NMR resonances derived from  $sp^2$ -hybridized carbon (commonly  $\delta_C > 110$  ppm) are expected to suffer more from (HSQC and HMBC) cross peak attenuation at higher F1 increments in 2D NMR spectra than other NMR resonances derived from  $sp^3$ -hybridized carbon (commonly  $\delta_C < 110$  ppm). Another, fundamentally different mechanism for enhanced longitudinal and transverse NMR relaxation is caused by free electrons in radicals and (paramagnetic) metal ions (Smernik and Oades, 1999, 2000, 2002). Metal ions in marine waters might coordinate several DOM ligand molecules (Hertkorn et al., 2004; Pohl et al., 2011; Vraspir and Butler, 2009; Witter et al., 2000) and therefore reduce their apparent mobility. In addition, coordination compounds of marine DOM with paramagnetic 3d metal ions like V, Mn, Cr, Fe, Co, Ni, and Cu will be plausible candidates for fast relaxing species of low NMR visibility (Bertini et al., 2008; see also section 3.14).

COSY and JRES NMR spectra are commonly computed with strongly truncating apodization functions which are designed to emphasize cross peak resolution and good pattern depiction (Cavanagh et al., 2007; Hertkorn et al., 2005). This behaviour will discriminate in favour of molecules which exceed a certain minimum abundance, a typical feature expected for common biological and abundant biogeochemical signatures. JRES NMR spectra offer rarely any liberty in selecting apodization functions beyond unshifted sine-bells in absolute value modulus to accommodate for complex cross peak amplitudes from remote  ${}^nJ_{HH}$  couplings and limited nominal NMR matrix size (Table S2). COSY NMR spectra with a  $\sim 300$  times larger nominal matrix size than available in JRES NMR spectra offer improved depiction of NOM organic complexity and liberty to employ mixed apodization functions to selectively enhance either cross peak resolution or S/N ratio. However, the antiphase character of COSY cross peaks will make them susceptible to attenuation and possibly self-cancellation in case of larger linewidths as a consequence of fast transverse NMR relaxation (Ernst et al., 1987). TOCSY NMR spectra offer absorptive lineshape cross peaks which are best suited for sensitive detection of minor NOM signatures, e.g. in the aromatic region, and offer  $1/2 \times {}^nJ_{HH}$  transfer time as opposed to  $1/{}^nJ_{HH}$  in COSY NMR spectra (Hertkorn et al., 2002b).

One-dimensional NMR spectra are commonly apodized to enhance S/N ratio by easily adjustable exponential multiplication which leaves spectral line area ratios invariant but not amplitude ratios. Hence, the matched filter concept (Spencer, 2010; Traficante and Rajabzadeh, 2000) is not unambiguously defined in marine SPE-DOM because of variable linewidth. The dramatic effects of educated apodization (Hertkorn et al., 2002a; Hoch & Stern, 1996) are demonstrated in Fig. S2 in which a TOCSY NMR spectrum of SPE-DOM 48 m is computed with an apodization designed to enhance S/N ratio at the expense of resolution (Fig. S2, panels 1A, 1B), whereas apodization optimized for resolution enhancement is applied in Fig. S2, panels 1C and 1D, respectively. Faint and rather expansive TOCSY cross peaks ( $\delta_H \sim 5 \dots 10$  ppm) became visible at large S/N ratio (Fig. S2, panel 1A) which were not at all discernible in case of resolution enhancement used (Fig. 2, panel 1C). Instead, resolution enhancement brought out sharp resolved cross peak patterns indicative of small, and relatively abundant molecules which otherwise were lost in expansive low-resolution cross peaks with minimal recognizable features.

NMR experiment	nominal NMR matrix size [Hz <sup>n</sup> ] (apparent pixel / voxel resolution at B <sub>0</sub> = 18.8 T)	intrinsic nominal sensitivity* (γ <sub>ex</sub> γ <sub>det</sub> <sup>3/2</sup> )	transfer mechanism and nominal time delay for magnetization transfer	F2/F1-apodization characteristics	general characteristics for the analysis of marine DOM
1D <sup>1</sup> H	10 × 800 = 8 × 10 <sup>3</sup>	1	10 μs (spin echo delay)	exponential (resolution)	sensitive, quantitative, strong signal overlap
1D <sup>13</sup> C	235 × 200 = 4.7 × 10 <sup>4</sup>	1.7 × 10 <sup>-4</sup>	10 μs (spin echo delay)	exponential (S/N ratio)	insensitive, informative, strong signal overlap
2D <sup>1</sup> H, <sup>1</sup> H JRES	8000 × 25 = 2 × 10 <sup>5</sup>	1	<sup>n</sup> J <sub>HH</sub> 100 ms	sine-bell (resolution)	good to reveal abundant molecular signatures; strong attenuation in case of fast transverse relaxation because of long duration of F1 increments
2D <sup>1</sup> H, <sup>1</sup> H COSY	8000 <sup>2</sup> = 6.4 × 10 <sup>7</sup>	1	1 / <sup>2-3</sup> J <sub>HH</sub> 133 ms	sine-bell (resolution) and exponential / sine (S/N ratio and resolution)	discriminates in favour of molecular signatures; transfer amplitude ~ 1/J <sub>HH</sub> ; improved resolution of smaller couplings J <sub>HH</sub> at higher number of F1 increments with attenuation from differential relaxation; cancellation of antiphase COSY cross peaks at higher linewidth is possible
2D <sup>1</sup> H, <sup>1</sup> H TOCSY	8000 <sup>2</sup> = 6.4 × 10 <sup>7</sup>	1	1 / 2 × <sup>2-5</sup> J <sub>HH</sub> 100 ms	exponential / sine (both)	sensitive for detection of minor signatures; absorptive lineshape: contributions from small resonances are fairly visible; transfer amplitude ~ 1/2 × J <sub>HH</sub>
2D <sup>1</sup> H, <sup>13</sup> C HSQC Family	8000 × 200 <sup>2</sup> = 3.2 × 10 <sup>8</sup>	2.5 × 10 <sup>-3</sup>	1 / <sup>1</sup> J <sub>CH</sub> 3.5 ms	exponential / sine (both)	absorptive lineshape; good combination of sensitivity and large information content; spectral editing feasible
2D <sup>1</sup> H, <sup>13</sup> C HMBC	8000 × 200 × 235 = 3.8 × 10 <sup>8</sup>	2.5 × 10 <sup>-3</sup>	1 / <sup>2-4</sup> J <sub>CH</sub> 150 ms	exponential / sine (S/N ratio)	discriminates in favour of abundant molecular signatures because of low sensitivity; excellent peak dispersion; allows assembly of extended spin systems across heteroatoms and quaternary carbon
2D <sup>13</sup> C, <sup>13</sup> C INADEQUATE	47000 <sup>2</sup> = 2.3 × 10 <sup>9</sup>	3.1 × 10 <sup>-6</sup>	1 / <sup>1</sup> J <sub>CC</sub> 25 ms	exponential / sine (S/N ratio)	ultimate information about CC connectivities; not practical for marine NOM at present because of limited sensitivity
3D <sup>1</sup> H, <sup>13</sup> C, <sup>1</sup> H HMQC-TOCSY	8000 <sup>2</sup> × 200 <sup>2</sup> = 2.6 × 10 <sup>12</sup>	2.5 × 10 <sup>-3</sup>	1 / 2 × <sup>2-5</sup> J <sub>HH</sub> + 1 / <sup>1</sup> J <sub>CH</sub> 100 ms	exponential / sine (S/N ratio)	excellent resolution with appreciable S/N ratio is now within realistic reach for high field NMR; in practice lesser utilizable resolution for <sup>13</sup> C (e.g. 1.7 × 10 <sup>7</sup> volumetric pixels; Simpson et al., 2003)

**Table S2.** Fundamental characteristics of key solution NMR experiments in the analysis of DOM (nominal bandwidth for <sup>1</sup>H NMR spectroscopy at B<sub>0</sub> = 18.8 T: 10 ppm × 800 Hz/ppm = 8000 Hz; <sup>13</sup>C NMR: 235 ppm × 200 Hz/ppm = 47 kHz nominal spectral width); \* intrinsic nominal sensitivity refers to natural isotopic abundance (<sup>1</sup>H: 99.985 %; <sup>13</sup>C: 1.1 %) and full relaxation.



**Fig. S1.** TOCSY NMR spectra of surface marine SPE-DOM (48 m) with (A, B) sensitivity enhanced apodization (exponential multiplication: EM = 7.5 Hz in F2;  $\pi/2.5$  shifted sine bell in F1) to emphasize (A) faint and extended NMR signatures of unsaturated chemical environments in marine DOM (cf. Fig. 12); (B) here, the aliphatic section is characterized by extensive overlap of abundant TOCSY cross peaks with limited overall resolution. (C, D) resolution enhanced apodization (Gaussian multiplication: EM: -0.4 Hz; GB: 0.6 Hz in F2;  $\pi/6$  shifted sine bell in F1) to emphasize depiction of (bio)molecular complexity in marine DOM. (C) cross peaks of protons attached to  $sp^2$ -hybridized carbon with resolved signatures of relatively abundant molecules (olefins and aromatics); (D) cross peaks of protons attached to  $sp^3$ -hybridized carbon (commonly  $\underline{H}CC$  and  $\underline{H}CO$  chemically environments) with resolved cross peak patterns for, e. g. alkylated / desoxy (dotted box a) and standard carbohydrates (dotted box b).

### SI 3 Previous $^1\text{H}$ and $^{13}\text{C}$ NMR studies of marine DOM

NMR spectroscopy of marine NOM has been regularly reviewed (Cook, 2004; Mopper et al., 2007; Simpson et al., 2011); here, this concise outline mainly refers to fundamental findings and concepts which initiated the current study. Early studies of marine organic matter almost exclusively focused on solid state  $^{13}\text{C}$  NMR spectroscopy of XAD isolates and tangential flow ultrafiltrates, and several key NOM building blocks were proposed such as peptides/proteins, carbohydrates, carboxylic acids and large proportion of common aliphatics. Meanwhile, solid state  $^{13}\text{C}$  NMR spectroscopy has produced fairly advanced spectral editing schemes for elucidating connectivities and functional groups in organic matter which are based on dipolar couplings and perceived atom mobilities rather than J-couplings as used in solution NMR spectral editing (Mao et al. 2011). A recent solid-state  $^{13}\text{C}$  NMR study has revealed heteropolysaccharides, peptides or amino sugars and COO-bonded alkyls as major constituents of marine ROED-DOM, in line with initial assessments but with considerably improved connectivity information (Mao, et al. 2012). The method of marine NOM isolation decisively affects the appearance of  $^{13}\text{C}$  NMR spectra of marine DOM; for example, average  $^{13}\text{C}$  NMR spectra of marine ROED-DOM closely correspond to a 45:55 mixture of UF-DOM (with enlarged carbohydrate content) and XAD isolates (increased aliphaticity and carboxylation) (Koprivnjak, et al. 2009).  $^{13}\text{C}$  NMR spectra of *riverine* C18-SPE-DOM revealed amides and oxygenated aliphatics such as aliphatic esters and ethers as well as hydroxylated aliphatics and abundant aliphatics (Kaiser et al., 2003). The authors are not aware of a dedicated  $^{13}\text{C}$  NMR study referring to the analysis of marine SPE-DOM.

Proton NMR studies of marine DOM show improved sensitivity but lesser structural resolving power than  $^{13}\text{C}$  NMR spectra and are still rather scarce. Main marine NOM constituents observed in  $^1\text{H}$  NMR spectra included prominent aliphatic lipids including peptide side chains, CRAM, acetates, a large diversity of carbohydrate units, and widely variable aromatics (Aluwihare et al., 2002; Repeta et al., 2002; Hertkorn et al., 2006).  $^1\text{H}$  NMR spectra revealed a strong accordance of marine DOM obtained by identical methods of isolation: 500 MHz NMR spectra of marine UF-DOM obtained from the Indian surface ocean and the equatorial Pacific surface ocean appeared nearly indistinguishable (Hertkorn et al., 2006; Repeta et al., 2002). A comparative assessment of SPE of marine waters using various adsorbent materials has depicted  $^1\text{H}$  NMR spectra of PPL-SPE-DOM and C18-SPE-DOM. In fact, the differences in between these two  $^1\text{H}$  NMR spectra were minor and in both cases the ratio of aromatic protons ( $\text{C}_{\text{ar}}\text{H}$ ), oxygenated ( $\text{OCH}$ ) and pure aliphatics ( $\text{CCH}$ ) ranged near 1:2:7 (Dittmar et al., 2008).

The powerful combination of mass spectrometry and NMR spectroscopy in the structural analysis of marine NOM has revealed carboxyl-rich alicyclic molecules (CRAM), itself a complex mixture of compounds as major constituents of marine UF-DOM, especially at larger depth in which CRAM comprised about half of the carbon of UF-DOM (Hertkorn et al., 2006). Combined NMR/FTMS characterization also has furnished improved assessment of chemical environments within atmospheric (Schmitt-Kopplin et al., 2010b), groundwater (Einsiedl et al., 2007) and extraterrestrial (Schmitt-Kopplin et al., 2010a) organic matter.



## SI 4 Previous FTICR mass spectrometry studies of marine DOM

Common high field FTICR mass spectra of marine DOM regularly show several thousands of resolved mass peaks of which from one third to one half can be attributed defined molecular compositions by virtue of excellent mass accuracy (D'Andrilli et al. 2010; Hertkorn et al. 2006; Koch et al. 2005; Koch and Dittmar 2006; Mopper et al. 2007). Most of FTMS studies to this day have used negative electrospray ionization (ESI) which preferentially ionizes water soluble hydrophilic compounds of marine DOM. Atmospheric photoionization (APPI) has resulted in improved coverage of aromatic and oxygen-deficient marine DOM molecules (D'Andrilli et al. 2010), reinforcing previous assessments that a realistic MS characterization of any NOM has to rely on data from several complementary ionization modes (Hertkorn et al. 2008). In both ESI and APPI studies of marine DOM composition, C<sub>n</sub>H<sub>m</sub>O<sub>q</sub>-molecular series have regularly comprised the majority of assigned molecular compositions (D'Andrilli et al. 2010; Hertkorn et al. 2006; Koch et al. 2005). In case of definitive absence of heteroatoms, FTICR mass spectra acquired at  $B_0 \geq 9.4$  T would produce unambiguous attribution of molecular compositions C<sub>n</sub>H<sub>m</sub>O<sub>q</sub> up to  $m/z \sim 928.2$  (Hertkorn, et al. 2008). However, challenges and ambiguities in the assignment of MS-derived marine DOM compositions arise from intrinsic uncertainty of mass peak definition from spectrometer imperfections (Koch and Dittmar 2006; Koch et al. 2007; Soule et al. 2010) and mass peak overlap and the supposable presence of heteroatoms in molecular compositions which occur in more numerous mathematical combinations than manageable as a whole (Kujawinski and Behn, 2006).

Improved reliability of marine DOM formula definition is available from elaborate data evaluation using the compositional space of chemical compositions (Hertkorn, et al. 2008) and a certain chemical plausibility for definition of unequivocal exclusion criteria (Koch and Dittmar 2006). Further improvements resulted from a priori and a posteriori exclusion of certain elements (Koch et al. 2007) together with extrapolation within molecular series of predefined building blocks, including stable isotopes such as <sup>13</sup>C and <sup>34</sup>S. Compositional network analysis, a recently introduced n-dimensional conceptual variant of Kendrick mass analysis, which defines mathematical relationships between all mass peaks in a FTICR mass spectrum, has resulted in more exhaustive and reliable attribution of molecular compositions in NOM (Tziotis et al., 2011). As a consequence, more numerous heteroatoms are now being identified in viable MS-derived marine NOM compositions.

Expansive data visualization schemes such as van Krevelen diagrams and Kendrick mass analyses are indispensable to assess MS-derived compositional characteristics of NOM (Hertkorn et al. 2008). Here, positioning within H/C against O/C van Krevelen diagrams according to structural relationships becomes more ambiguous in presence of heteroatoms (e.g. S and P might carry oxygen in elevated oxidation states; nitrogen carries hydrogen into molecular compositions: Kroll et al., 2011; Schmitt-Kopplin et al., 2010b). In case of larger molecules, extremely numerous isomers will likely project onto single molecular compositions (Hertkorn et al. 2007), further obliterating relationships between composition and perceived structure. MS-based proposals of chemical structures of marine DOM molecules appear reasonable only for either extremely hydrogen deficient (Koch and Dittmar 2006) or fully saturated molecules; relaxed requirements apply for perceived biosignatures in marine DOM for which independent information (like metabolic pathway maps) are available for comparison (Longnecker and Kujawinski, 2011).

Notably, the classical depiction of mass spectra (mass against intensity) fails to properly evaluate the intrinsic molecular complexity of NOM because it solely refers to main fragments with mass differences exceeding that of double bond equivalents (DBE,  $m/z = 2.1057$  Da) (Hertkorn, et al. 2008). Chemically similar NOM molecules

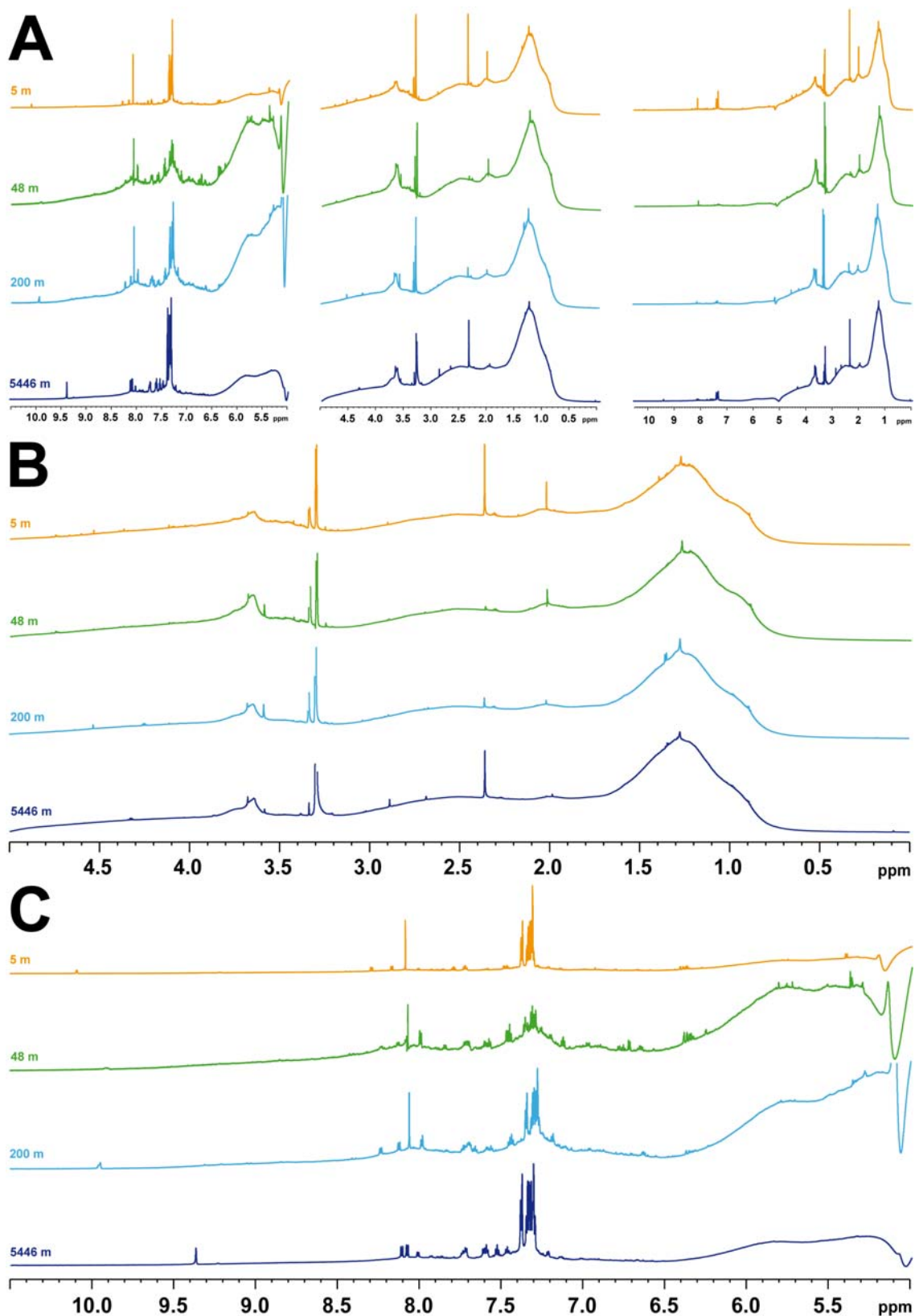
such as CH<sub>2</sub> homologues ( $\Delta m = 14.0156$  Da) will be always widely separated in FTICR mass spectra. Remarkably, the mandatory compositional (and structural) dissimilarity and mass spacing  $\Delta m$  between pairs of mass peaks in clusters of nominal mass are inversely related; this refers to any composition C<sub>n</sub>H<sub>m</sub>O<sub>q</sub>Z<sub>Σ</sub> (Z<sub>Σ</sub>: any combination of elements: Hertkorn, et al. 2008). Hence, slices across single nominal mass clusters are much better suited to depict compositional variance and molecular diversity of NOM.

**Table S3.** Acquisition parameters of NMR spectra, shown according to figures. PK: probeheads used for acquisition of NMR spectra ( $B_0 = 18.8$  T), QCI: cryogenic inverse geometry 5 mm z-gradient <sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N/<sup>31</sup>P QCI probe; QCO: cryogenic classical geometry 3 mm z-gradient <sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N/<sup>31</sup>P probe; 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/GM: line broadening factor [Hz]; QS: shifted square sine bell; SI: sine bell); PR1, PR2: coefficients used for windowing functions WDW1, WDW2, EM/GM are given in [Hz], SI/QS derived functions indicate shift by  $\pi/n$ .

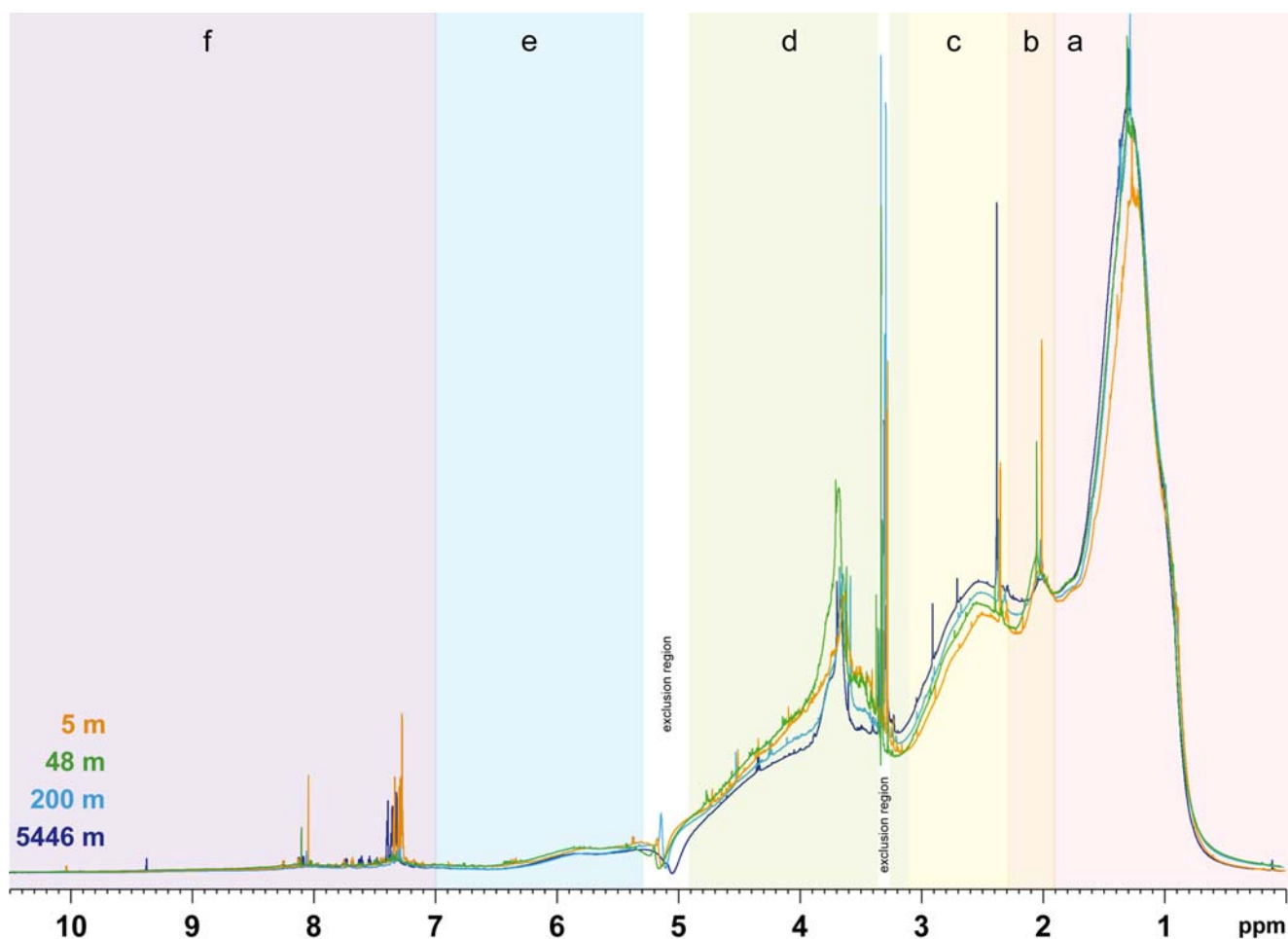
spectrum	Figure	PK	NS	AQ [ms]	D1 [ms]	NE	WDW1	WDW2	PR1	PR2
<sup>1</sup> H, <sup>1</sup> H TOCSY	S1A, 1B	QCI	16	750	3250	1600	SI	EM	2.5	7.5
<sup>1</sup> H, <sup>1</sup> H TOCSY	S1C, 1D	QCI	16	750	3250	1600	SI	GM	6	-0.4 / 0.6
<sup>1</sup> H NMR	2	QCI	512	5000	10000	-	-	EM	-	1
<sup>13</sup> C NMR	3: 5 m	QCO	36064	1000	19000	-	-	EM	-	12.5 / 5 / 1
<sup>13</sup> C NMR	3: 48 m	QCO	22432	1000	19000	-	-	EM	-	12.5 / 5 / 1
<sup>13</sup> C NMR	3: 200 m	QCO	21146	1000	19000	-	-	EM	-	12.5 / 5 / 1
<sup>13</sup> C NMR	3: 5446 m	QCO	10640	1000	19000	-	-	EM	-	12.5 / 5 / 1
<sup>13</sup> C DEPT	4A : FISH A	QCI	16384	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4A : FISH B	QCI	32768	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4A : DEPT-135	QCI	16384	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4B : FMAX A	QCI	14400	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4B : FMAX B	QCI	28800	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4B : DEPT-135	QCI	14400	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4C : 200 A	QCI	17920	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4C : 200 B	QCI	35840	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4C : DEPT-135	QCI	17920	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4D : 5446 A	QCI	25632	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4D : 5446 B	QCI	23937	1000	2000	-	-	EM	-	12.5
<sup>13</sup> C DEPT	4D : DEPT-135	QCI	25345	1000	2000	-	-	EM	-	12.5
<sup>1</sup> H, <sup>1</sup> H JRES	5A, 5B	QCI	80	750	750	80	QS	QS	0	0
<sup>1</sup> H, <sup>1</sup> H JRES	5C, 5D	QCI	128	750	750	80	QS	QS	0	0
<sup>1</sup> H, <sup>1</sup> H JRES	5E, 5F	QCI	80	750	750	80	QS	QS	0	0
<sup>1</sup> H, <sup>1</sup> H JRES	5G, 5H	QCI	128	750	750	80	QS	QS	0	0
<sup>1</sup> H, <sup>1</sup> H COSY	6A, 6B	QCI	16	750	2250	1600	QS	EM	3	2
<sup>1</sup> H, <sup>1</sup> H COSY	6C, 6D	QCI	64	750	2250	1600	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>1</sup> H COSY	6E, 6F	QCI	16	750	2250	1600	QS	EM	3	2
<sup>1</sup> H, <sup>1</sup> H COSY	6G, 6H	QCI	64	750	2250	1600	QS	EM	2.5	2.5

spectrum	Figure	PK	NS	AQ [ms]	D1 [ms]	NE	WDW1	WDW2	PR1	PR2
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>3</sub> : 7A	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>3</sub> : 7A	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>3</sub> : 7B	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>3</sub> : 7C	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>3</sub> : 7D	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>2</sub> : 7E	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>2</sub> : 7F	QCI	64	250	1250	1024	QS	EM	2.5	2.5
spectrum	Figure	PK	NS	AQ [ms]	D1 [ms]	NE	WDW1	WDW2	PR1	PR2
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>2</sub> : 7E	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>2</sub> : 7F	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>2</sub> : 7G	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>2</sub> : 7H	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>3</sub> : 8A,8C	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C DEPT- HSQC	CH <sub>2</sub> : 8B	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C HSQC- TOCSY	9A	QCI	288	250	1250	394	QS	EM	2	2
<sup>1</sup> H, <sup>13</sup> C HSQC- TOCSY	9B, 9E	QCI	320	250	1250	400	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C HSQC- TOCSY	9C	QCI	64	250	1250	1024	QS	EM	2.5	2.5
<sup>1</sup> H, <sup>13</sup> C HSQC- TOCSY	9D	QCI	320	250	1250	400	QS	EM	2.5	12.5
<sup>1</sup> H, <sup>13</sup> C HSQC- TOCSY	9D	QCI	320	250	1250	400	QS	EM	2.5	12.5
<sup>1</sup> H, <sup>13</sup> C HMBC	10A, 10B	QCI	4000	250	1250	117	QS	EM	2.5	5
<sup>1</sup> H, <sup>13</sup> C HMBC	10C, 10D	QCI	4000	750	750	161	QS	EM	2.5	5
<sup>1</sup> H, <sup>1</sup> H TOCSY	12A, 12E	QCI	16	750	3250	1600	SI	EM	2.5	7.5
<sup>1</sup> H, <sup>1</sup> H TOCSY	12B, 12F	QCI	32	750	3250	1600	SI	EM	2.5	7.5
<sup>1</sup> H, <sup>1</sup> H TOCSY	12C, 12G	QCI	8	750	3250	1600	SI	EM	2.5	7.5
<sup>1</sup> H, <sup>1</sup> H TOCSY	12D, 12H	QCI	32	750	3250	1600	SI	EM	2.5	7.5
<sup>1</sup> H, <sup>13</sup> C HSQC	13A, 13E, 13I	QCI	400	250	1250	137	QS	EM	2.5	5
<sup>1</sup> H, <sup>13</sup> C HSQC	13B, 13F	QCI	400	250	1250	142	QS	EM	2.5	5
<sup>1</sup> H, <sup>13</sup> C HSQC	13C, 13G	QCI	800	250	1250	137	QS	EM	2.5	5
<sup>1</sup> H, <sup>13</sup> C HSQC	13D, 13H	QCI	1800	250	1250	140	QS	EM	2.5	5

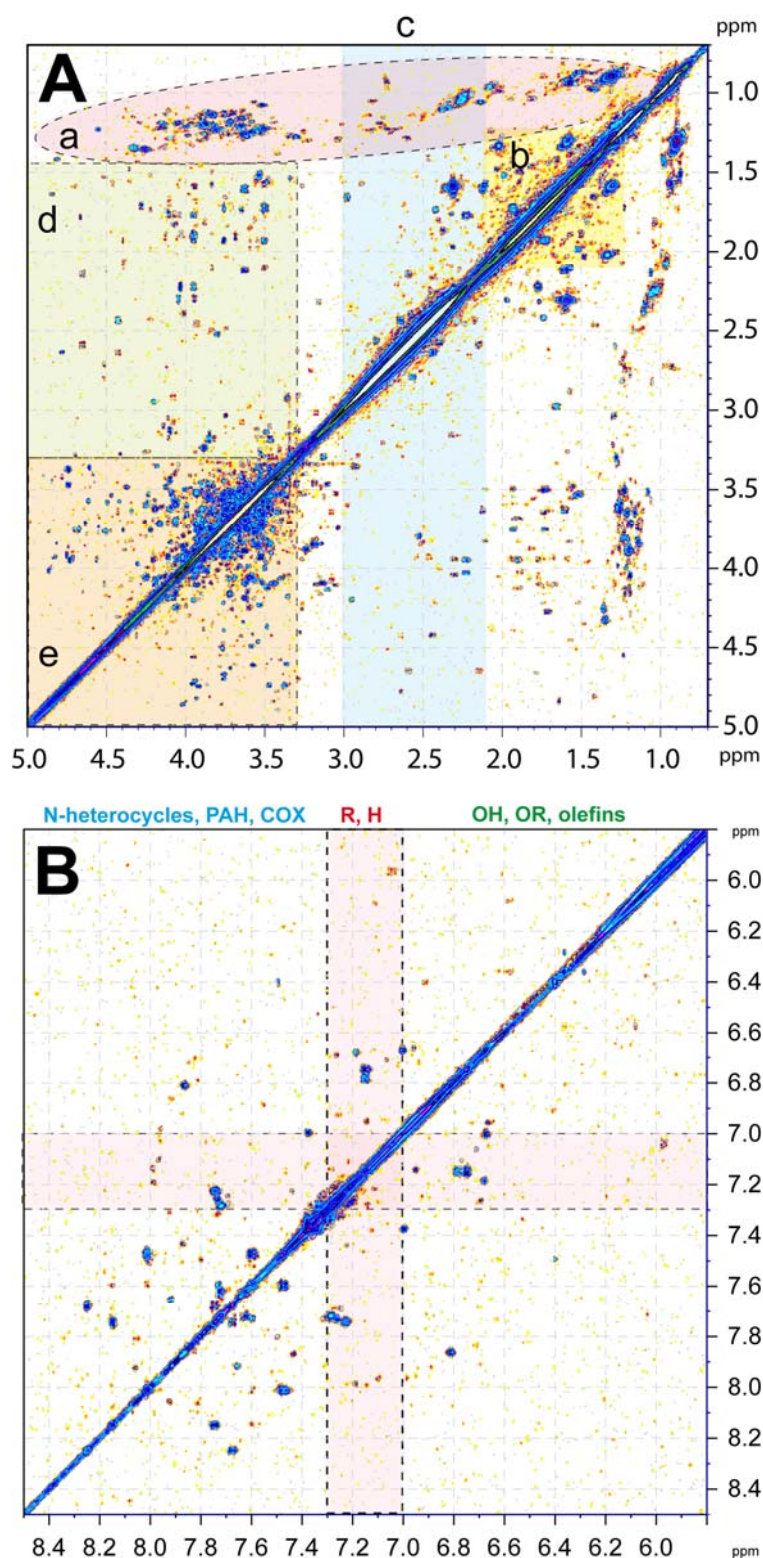
**Table S3.** Acquisition parameters of NMR spectra, shown according to figures. PK: probeheads used for acquisition of NMR spectra ( $B_0 = 18.8$  T), QCI: cryogenic inverse geometry 5 mm z-gradient <sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N/<sup>31</sup>P QCI probe; QCO: cryogenic classical geometry 3 mm z-gradient <sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N/<sup>31</sup>P probe; 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/GM: line broadening factor [Hz]; QS: shifted square sine bell; SI: sine bell); PR1, PR2: coefficients used for windowing functions WDW1, WDW2, EM/GM are given in [Hz], SI/QS derived functions indicate shift by  $\pi/n$ .



**Fig. S2.**  $^1\text{H}$  NMR spectra of marine SPE-DOM obtained by solid phase extraction (PPL) from 3.1°E; -17.7°S (Angola basin) at different depths according to color: orange: 5 m (near surface photic zone); green: 48 m (fluorescence maximum); blue: 200 m (upper mesopelagic zone); dark blue: 5446 m (30 m above ground) with respective spectral intensities scaled to 100% of maximum intensity (with methanol excluded) within the regions shown. (A): three spectral ranges,  $\delta_{\text{H}} = 5-10.5 / 0-5 / 0-10.5$  ppm; (B): aliphatic section,  $\delta_{\text{H}} = 0-5$  ppm. (C): unsaturated protons, bound to  $\text{sp}^2$ -hybridized carbon;  $\delta_{\text{H}} = 5-10.5$  ppm and anomeric protons ( $\delta_{\text{H}} < 5.5$  ppm); cf. Fig. 2.

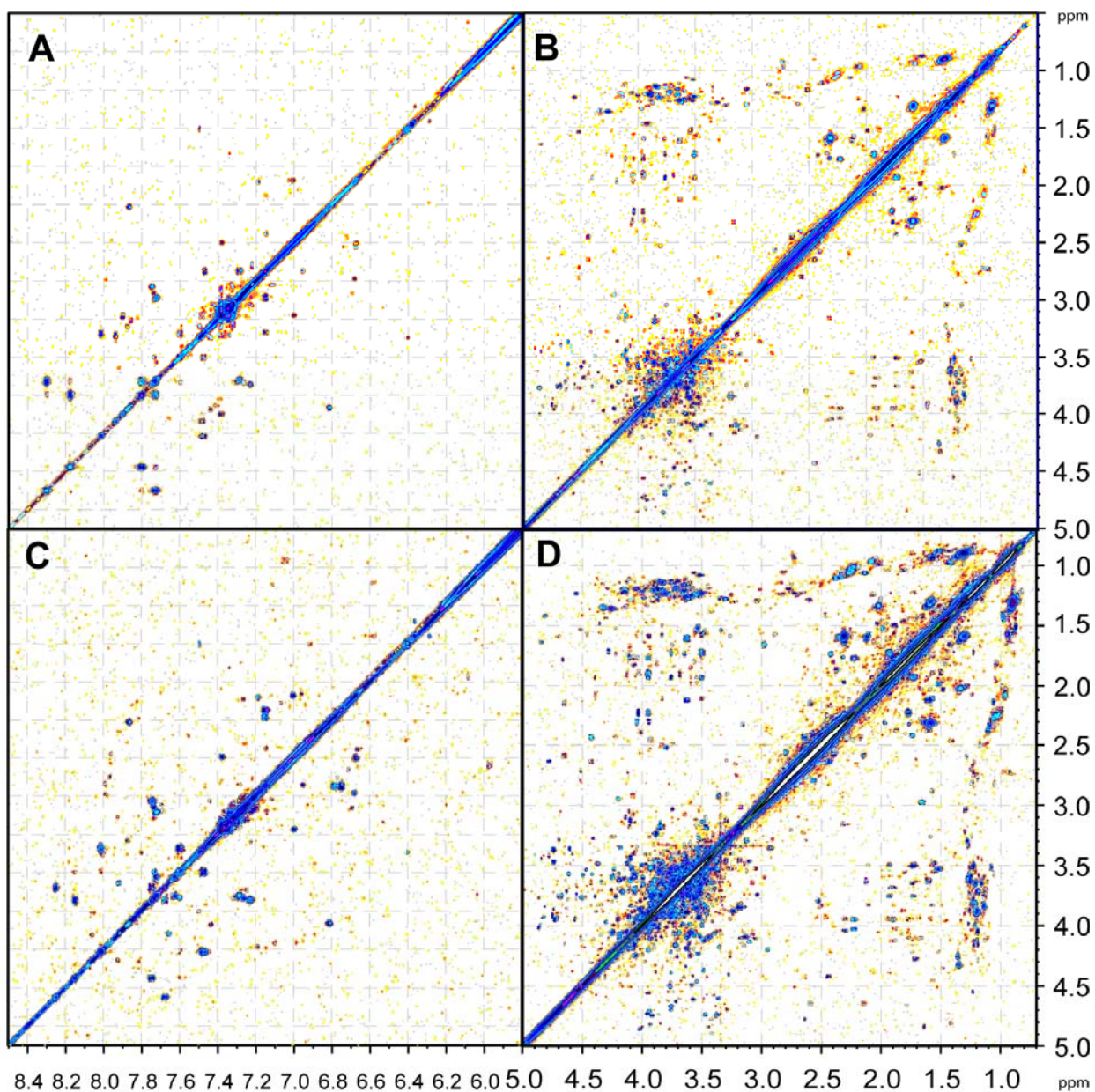


**Fig. S3.**  $^1\text{H}$  NMR spectra of four marine SPE-DOM acquired with solvent suppression and exclusion regions used in the computation of NMR section integrals and overlay NMR spectra (Fig. 2 and this figure) which denote  $\text{HD}_2\text{COD}$  and residual HDO. Intensities are normalized to 100% total integral in the entire chemical shift range shown ( $\delta_{\text{H}} = 0\text{-}10.5$  ppm). Fundamental substructures are indicated from higher to lower field (from right to left), (a) aliphatics,  $\text{HCCC}$ ; (b) “acetate-analogue”,  $\text{H}_3\text{CC}(=\text{O})\text{-O-}$ ; (c) carboxyl-rich alicyclic materials (CRAM),  $\text{HC}(\text{C})\text{-COX}$ ; (d) “carbohydrate-like” and methoxy,  $\text{HCO}$ ; (e) olefinic,  $\text{HC}=\text{C}$ ; and (f) aromatic NMR resonances  $\text{HC}_{\text{ar}}$ . Superimposed small NMR resonances indicative of comparatively abundant biological and biogeochemical molecules were most significant in the aromatic section (f), well noticeable in sections (d) and (e) and of continual lesser occurrence in the order  $c > b > a$ .

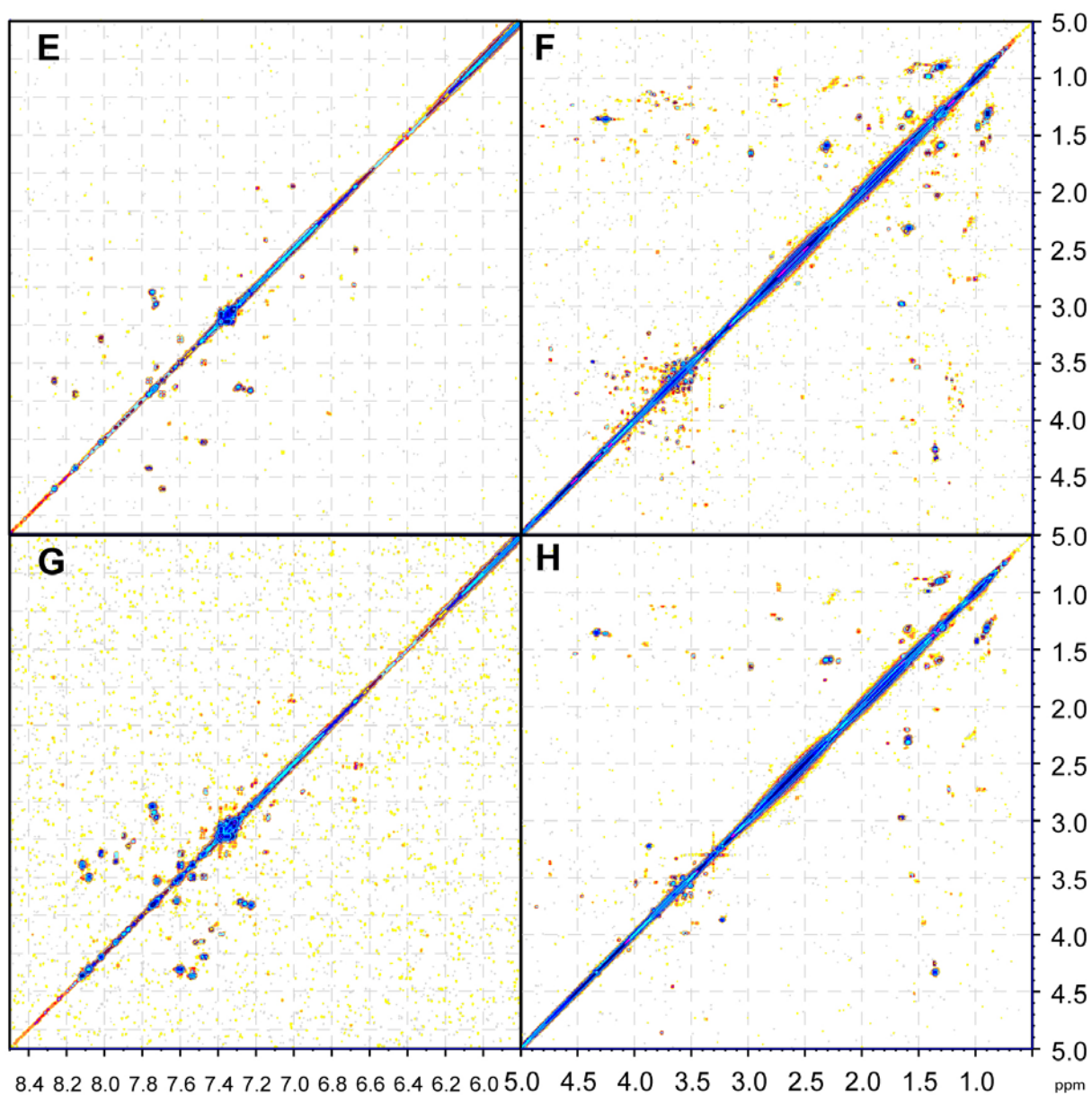


**Fig. S4.**  $^1\text{H}$ ,  $^1\text{H}$  COSY NMR spectrum of marine surface SPE-DOM 48m (cf. Fig. 6C, 6D) with key substructures indicated. Panel A: aliphatic section; section a:  $\text{H}_3\text{CCH}$  cross peaks related to chain terminating methyl groups, for  $\delta_{\text{H}} > 3$  ppm with oxygenated carbon ( $\text{H}_3\text{CCH-O-}$ ); section b: intra aliphatic  $\text{HC-CH-CH}$  cross peaks excluding methyl; section c:  $\text{HCH-COX}$  cross peaks with carbonyl derivatives; section d: functionalized aliphatics connected with oxygenated carbon  $\text{Z-C-HCH-O}$  [ $\text{Z} = \text{O (N)}$ ]; section e:  $-\text{O-HCH-O-}$  cross peaks, carbohydrates and esters, ethers and alcohols. Panel B: aromatic section with key substituents provided: electron withdrawing, polycyclic aromatic hydrocarbons and six-membered N-heterocycles ( $\delta_{\text{H}} > 7.3$  ppm), neutral ( $\delta_{\text{H}} \sim 7\text{-}7.3$  ppm), electron donating, olefins and five membered heterocycles ( $\delta_{\text{H}} < 7$  ppm); aromatic COSY cross peaks represent rather abundant molecular signatures of marine DOM.



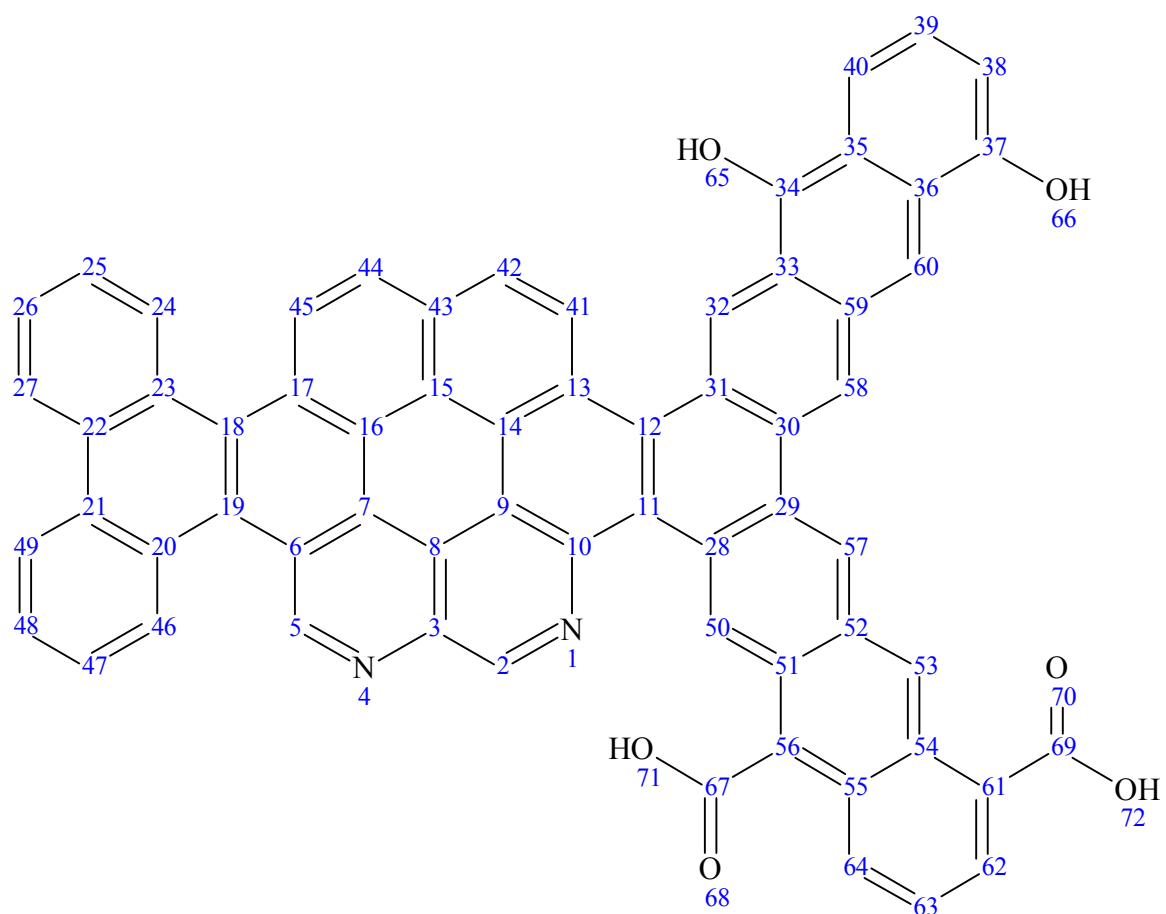


**Fig. S5A.**  $^1\text{H}$ ,  $^1\text{H}$  COSY NMR spectra of (A, C) downfield ( $\delta_{\text{H}} = 5.8\text{-}8.5$  ppm; aromatic and olefinic  $\text{C-C}_{\text{sp}^2}\text{H}$ - $\text{C}_{\text{sp}^2}\text{H-C}$  cross peaks) and (B, D) upfield  $^1\text{H}$  NMR chemical shift region ( $\delta_{\text{H}} = 0.5\text{-}5.0$  ppm; aliphatic  $\text{C-HC-HC-X}$  (X: C, N, O); A, B: 5 m (near surface photic zone); C, D: 48 m (fluorescence maximum)).



**Fig. S5B.**  $^1\text{H}$ ,  $^1\text{H}$  COSY NMR spectra of (E, G) downfield ( $\delta_{\text{H}} = 5.8\text{-}8.5$  ppm; aromatic and olefinic C-C<sub>sp<sup>2</sup></sub>-H-C<sub>sp<sup>2</sup></sub>-H-C cross peaks) and (F, H) upfield  $^1\text{H}$  NMR chemical shift region ( $\delta_{\text{H}} = 0.5\text{-}5.0$  ppm; aliphatic C-HC-HC-X (X: C, N, O); E, F: 200 m (upper mesopelagic zone); G, H: 5446 m (abyssopelagic, 30 m above seafloor)).

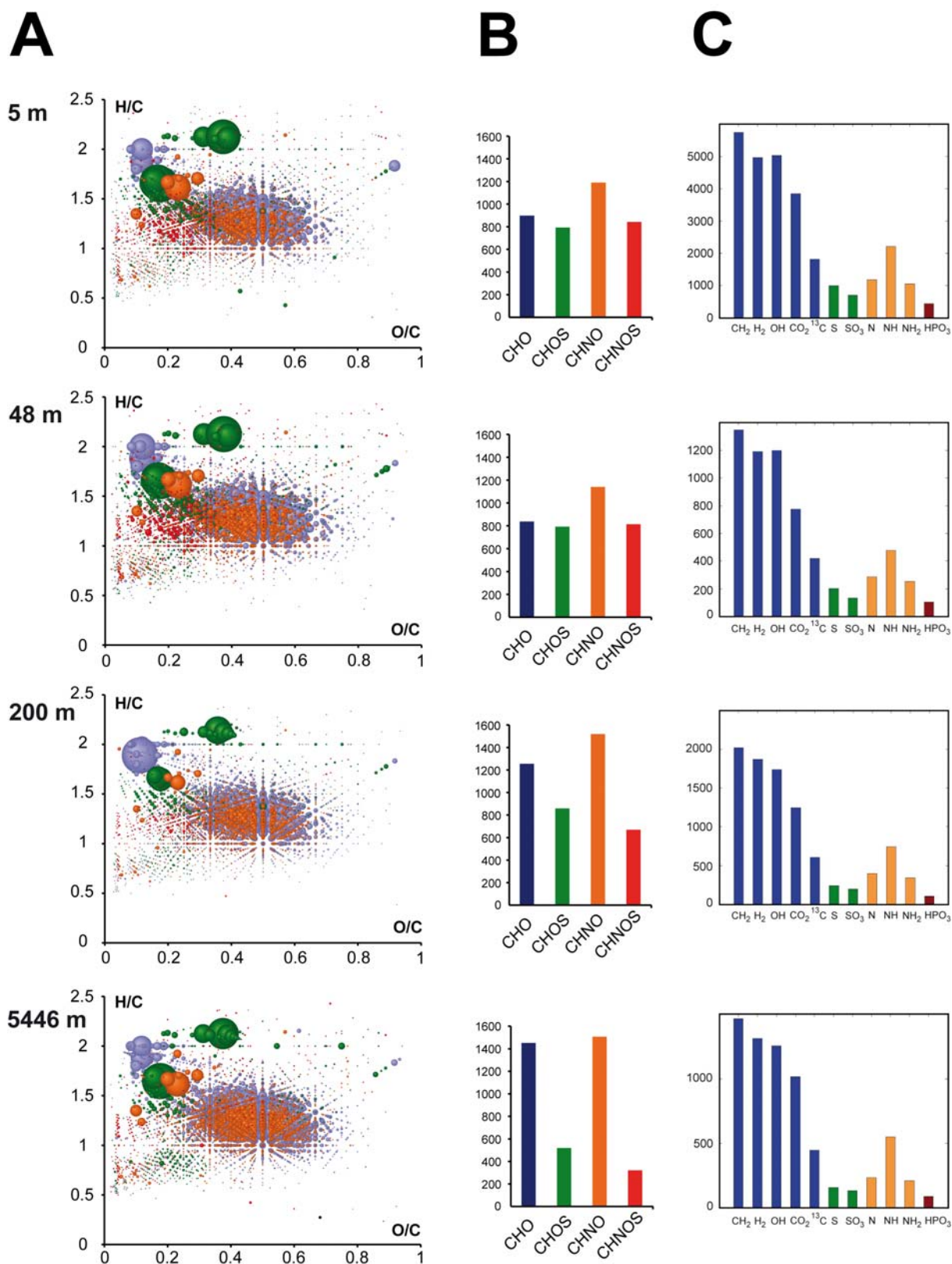




**Fig. S6.** ACD software proposed numbering scheme of PAH model molecule  $C_{65}H_{34}N_2O_6$  (IUPAC mass. 938.975 Da; cf. Fig. 11 and Table S4).

carbon number	CH <sub>n</sub>	δ( <sup>13</sup> C) [ppm]	δ( <sup>1</sup> H) [ppm]	carbon number	CH <sub>n</sub>	δ( <sup>13</sup> C) [ppm]	δ( <sup>1</sup> H) [ppm]	carbon number	CH <sub>n</sub>	δ( <sup>13</sup> C) [ppm]	δ( <sup>1</sup> H) [ppm]
2	CH	164.32	9.18	25	CH	126.55	7.75	47	CH	129.36	7.72
3	C	151.49		26	CH	126.44	7.61	48	CH	126.44	7.61
5	CH	154.56	9.90	27	CH	125.15	8.57	49	CH	126.78	8.66
6	C	130.92		28	C	140.71		50	CH	136.73	9.43
7	C	131.05		29	C	136.60		51	C	141.62	
8	C	132.93		30	C	147.56		52	C	137.65	
9	C	133.41		31	C	130.89		53	CH	121.83	8.61
10	C	141.35		32	CH	136.41	8.86	54	C	134.91	
11	C	133.85		33	C	127.98		55	C	134.91	
12	C	132.19		34	C	150.72		56	C	140.72	
13	C	135.02		35	C	128.97		57	CH	134.28	9.52
14	C	134.79		36	C	128.97		58	CH	125.79	8.83
15	C	148.67		37	C	154.47		59	C	139.28	
16	C	132.73		38	CH	109.77	7.04	60	CH	110.67	8.84
17	C	127.11		39	CH	123.76	7.14	61	C	134.90	
18	C	123.78		40	CH	123.22	7.94	62	CH	132.45	8.7
19	C	127.60		41	CH	131.74	9.26	63	CH	129.81	7.47
20	C	131.30		42	CH	132.58	8.47	64	CH	133.78	8.47
21	C	141.33		43	C	131.74		67	C	169.89	
22	C	131.79		44	CH	132.58	8.47	69	C	168.82	
23	C	138.69		45	CH	129.05	9.13				
24	CH	132.52	9.07	46	CH	132.12	9.14				

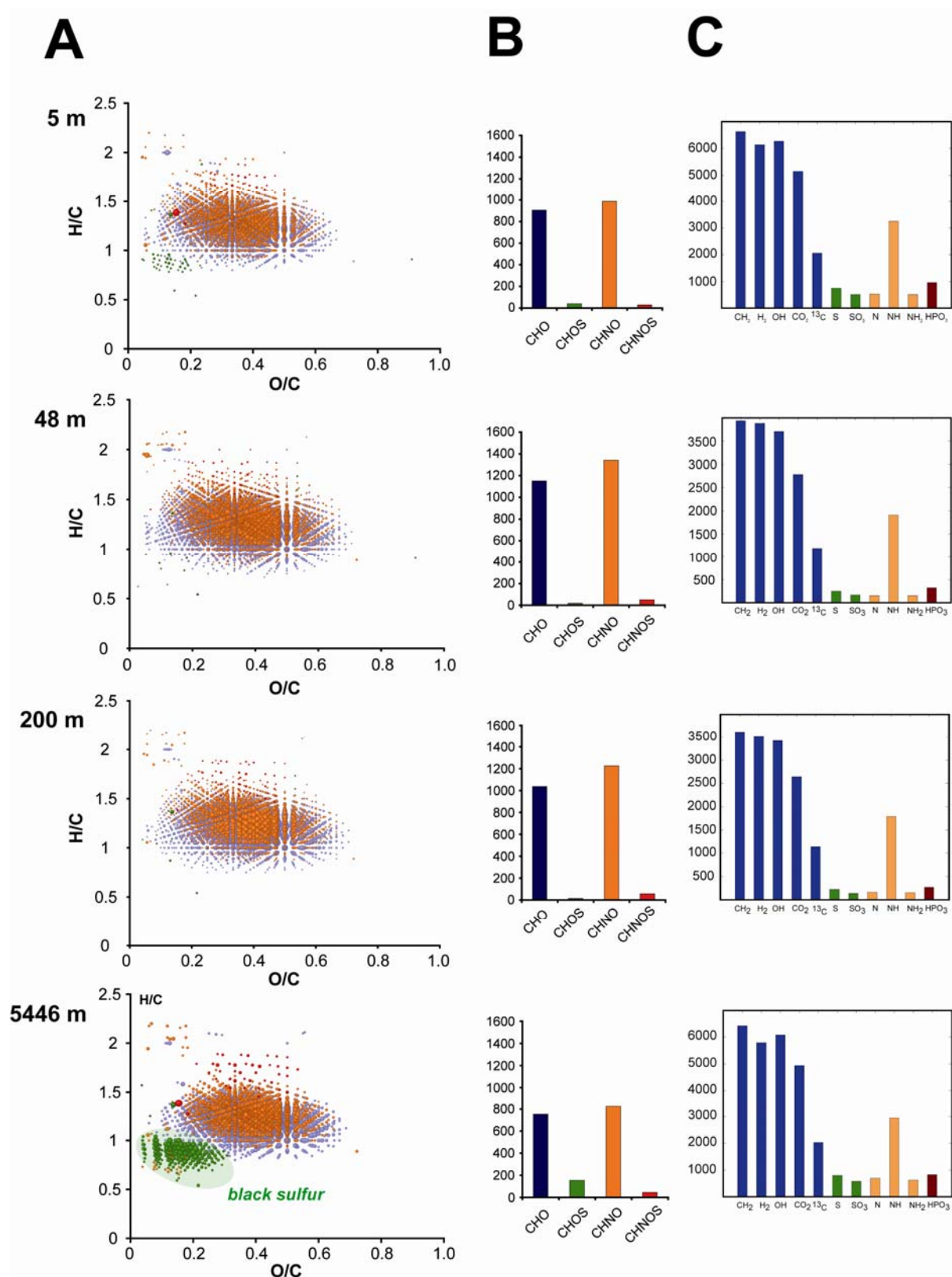
**Table S4.** computed proton and carbon NMR chemical shift of PAH model C<sub>65</sub>H<sub>34</sub>N<sub>2</sub>O<sub>6</sub> (IUPAC mass. 938.975 Da; cf. Fig. 11 and Fig. S5)



**Fig. S7.** Negative electrospray 12T FTICR mass spectra of marine SPE-DOM; from top to bottom: 5 m (near surface photic zone); 48 m (fluorescence maximum); 200 m (upper mesopelagic zone); 5446 m (30 m above ground). Panel A: H/C versus O/C van Krevelen diagrams of FTICR mass spectra from 150-700 Da with color code as provided in panel B: histograms represent counts of respective unique CHO, CHOS, CHNO and CHNOS containing molecules (cf. text). Panel C: histograms denoting frequency of major fragments in FTICR mass spectra (Schmitt-Kopplin et al., 2010a).

metal	FISH	FMAX	200 m	5446 m
Co	59	48	309	72
Cr	1626	1016	1760	1676
Cu	1407	4927	1359	2187
Fe	75305	7375	3057	7327
Mn	3080	126	1230	255
Ni	10339	7130	765	586
Al	29937	9884	13250	19731
As	12191	13678	13525	5740
B	42282	22644	7038	94615
Ba	870	1310	1637	4500
Ca	49380	44064	47124	136153
Cd	237	135	106	779
Hg	61	47	21	23
Mg	19042	11108	36414	42404
Mo	15092	8629	7252	8711
Na	351834	169831	172586	481727
P	270357	200432	138007	79903
Pb	1117	842	1178	172
S	1145004	813967	645666	435574
Sb	171	104	133	100
Sn	151	164	228	59
Sr	302	373	334	395
Th	45	21	20	12
Ti	512	334	1655	199
U	1673	805	456	372
Zn	10432	14443	10159	20365
Be	4	4	4	3
Bi	4	3	3	3
Cs	4	4	4	3
K	33332	32742	32436	28846
Li	722	710	707	626
Se	4290	4192	4162	3692
V	1710	1677	1671	1480

**Table S5.** Metal concentration (ng metal / L methanol extract) for four marine SPE-DOM (gray cells: metal concentrations were below detection limit; cf. Fig. 20).



**Fig. S8.** Positive electrospray 12T FTICR mass spectra of marine SPE-DOM; from top to bottom: 5 m (near surface photic zone); 48 m (fluorescence maximum); 200 m (upper mesopelagic zone); 5446 m (30 m above ground). Panel A: H/C versus O/C van Krevelen diagrams of FTICR mass spectra from 150-700 Da with color code as provided in panel B: histograms represent counts of respective unique CHO, CHOS, CHNO and CHNOS containing molecules (cf. text). Panel C: histograms denoting frequency of major fragments in FTICR mass spectra (Schmitt-Kopplin et al., 2010a).

## REFERENCES

- Aluwihare, L. I., Repeta, D. J., and Chen, R. F.: Chemical composition and cycling of dissolved organic matter in the mid-atlantic bight, *Deep-Sea Research, Part II: Topical Studies in Oceanography*, 49, 4421-4437, 2002.
- Amon, R. M. W., Benner, R.: 1996. Bacterial utilization of different size classes of dissolved organic matter. *Limnol. Oceanogr.*, 41, 41-51, 1996.
- Bakmutov, V., I.: *Practical NMR Relaxation for Chemists*, John Wiley & Sons, Chichester, England, 2004.
- Bertini, I., Luchinat, C., Parigi, G., and Pierattelli, R.: Perspectives in paramagnetic NMR of metalloproteins, *Dalton T*, 3782-3790, 2008.
- Buddrus, J., Burba, P., Lambert, J., and Herzog, H.: Quantification of partial structures of aquatic humic substances by one- and two-dimensional solution <sup>13</sup>C nuclear magnetic resonance spectroscopy, *Anal. Chem.*, 61, 628-631, 1989.
- Carper, W. R., Wahlbeck, P. G., and Dolle, A.: C-13 nmr relaxation rates: Separation of dipolar and chemical shift anisotropy effects, *J. Phys. Chem. A.*, 108, 6096-6099, 2004.
- Case, D. A.: The use of chemical shifts and their anisotropies in biomolecular structure determination, *Curr. Opin. Struc. Biol.*, 8, 624-630, 1998.
- Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., Rance, M., Skelton, N. J.: *Protein NMR Spectroscopy*, Elsevier, Burlington, USA, 2007.
- Cook, R. L., Langford, C. H., Yamdagni, R., and Preston, C. M.: A modified cross-polarization magic angle spinning c-13 nmr procedure for the study of humic materials, *Anal. Chem.*, 68, 3979-3986, 1996.
- Cook, R. L.: Coupling NMR to NOM. *Anal. Bioanal. Chem.*, 378, 1484-1503, 2004.
- D'Andrilli, J., Dittmar, T., Koch, B.P., Purcell, J.M., Marshall, A.G., Cooper, W.T.: Comprehensive characterization of marine dissolved organic matter by Fourier transform ion cyclotron resonance mass spectrometry with electrospray and atmospheric pressure photoionization. *Rapid Commun. Mass Sp.*, 24, 643-650, 2010.
- Dittmar, T., Koch, B., Hertkorn, N., and Kattner, G.: A simple and efficient method for the solid-phase extraction of dissolved organic matter (spe-DOM) from seawater, *Limnol. Oceanogr.-Meth.*, 6, 230-235, 2008.
- Effemey, M., Lang, J., and Kowalewski, J.: Multiple-field carbon-13 and proton relaxation in sucrose in viscous solution, *Magn. Reson. Chem.*, 38, 1012-1018, 2000.
- Einsiedl, F., Hertkorn, N., Wolf, M., Frommberger, M., Schmitt-Kopplin, P., and Koch, B. P.: Rapid biotic molecular transformation of fulvic acids in a karst aquifer, *Geochim. Cosmochim. Ac.*, 71, 5474-5482, 2007.
- Ernst, R. R., Bodenhausen, G., Wokaun, A.: *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford, 1987.
- Hansen, A. L., and Al-Hashimi, H. M.: Dynamics of large elongated RNA by NMR carbon relaxation, *J. Am. Chem. Soc.*, 129, 16072-16082, 2007.
- Hertkorn, N., Permin, A., Perminova, I., Kovalevskii, D., Yudov, M., Petrosyan, V., and Kettrup, A.: Comparative analysis of partial structures of a peat humic and fulvic acid using one- and two-dimensional nuclear magnetic resonance spectroscopy, *J. Environ. Qual.*, 31, 375-387, 2002.

- Hertkorn, N., Günzl, A., Freitag, D., Kettrup, A.: Nuclear Magnetic Resonance Spectroscopy Investigations of Silylated Refractory Organic Substances, in: *Refractory Organic Substances in the Environment (DFG-Schwerpunktprogramm Refraktäre Organische Säuren in Gewässern)*, edited by: Frimmel, F. H., Abbt-Braun, G., Wiley-VCH, p. 129-145, 2002.
- Hertkorn, N., Perdue, E. M., and Kettrup, A.: A potentiometric and  $^{113}\text{Cd}$  nmr study of cadmium complexation by natural organic matter at two different magnetic field strengths, *Anal. Chem.*, 76, 6327-6341, 2004.
- Hertkorn, N., and Kettrup, A.: Molecular level structural analysis of natural organic matter and of humic substances by multinuclear and higher dimensional NMR spectroscopy, in: *Use of humic substances to remediate polluted environments: From theory to practice*, edited by: Perminova, I., Hatfield, K., and Hertkorn, N., Springer Netherlands, 391-435, 2005.
- Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., Kettrup, A., and Hedges, J. I.: Characterization of a major refractory component of marine dissolved organic matter, *Geochim. Cosmochim. Ac.*, 70, 2990-3010, 2006.
- Hertkorn, N., Ruecker, C., Meringer, M., Gugisch, R., Frommberger, M., Perdue, E. M., Witt, M., and Schmitt-Kopplin, P.: High-precision frequency measurements: Indispensable tools at the core of the molecular-level analysis of complex systems, *Analytical and Bioanalytical Chemistry*, 389, 1311-1327, 2007.
- Hertkorn, N., Frommberger, M., Witt, M., Koch, B. P., Schmitt-Kopplin, P., and Perdue, E. M.: Natural organic matter and the event horizon of mass spectrometry, *Anal. Chem.*, 80, 8908-8919, 2008.
- Hoch, J., C., Stern, A., S.: *NMR Data Processing*, John Wiley & Sons, Wiley-Liss, New York, USA, 1996.
- Kaiser, E., Simpson, A. J., Dria, K. J., Sulzberger, B., and Hatcher, P. G.: Solid-state and multidimensional solution-state NMR of solid phase extracted and ultrafiltered riverine dissolved organic matter, *Environ. Sci. Technol.*, 37, 2929-2935, 2003.
- Keeler, J.: *Understanding NMR Spectroscopy*, John Wiley & Sons, Chichester, 2005.
- Koch, B. P., Witt, M. R., Engbrodt, R., Dittmar, T., and Kattner, G.: Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization fourier transform ion cyclotron resonance mass spectrometry, *Geochim. Cosmochim. Ac.*, 69, 3299-3308, 2005.
- Koch, B. P., and Dittmar, T.: From mass to structure: An aromaticity index for high-resolution mass data of natural organic matter, *Rapid. Commun. Mass. Sp.*, 20, 926-932, 2006.
- Koch, B. P., Dittmar, T., Witt, M., Kattner, G.: Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter. *Anal. Chem.*, 79, 1758-1763, 2007.
- Koprivnjak, J. F., Pfromm, P. H., Ingall, E., Vetter, T. A., Schmitt-Kopplin, P., Hertkorn, N., Frommberger, M., Knicker, H., and Perdue, E. M.: Chemical and spectroscopic characterization of marine dissolved organic matter isolated using coupled reverse osmosis-electrodialysis, *Geochim. Cosmochim. Ac.*, 73, 4215-4231, 2009.
- Korzhnev, D. M., Kloiber, K., and Kay, L. E.: Multiple-quantum relaxation dispersion NMR spectroscopy probing millisecond time-scale dynamics in proteins: Theory and application, *J. Am. Chem. Soc.*, 126, 7320-7329, 2004a.
- Korzhnev, D. M., Salvatella, X., Vendruscolo, M., Di Nardo, A. A., Davidson, A. R., Dobson, C. M., and Kay, L. E.: Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR, *Nature*, 430, 586-590, 2004b.

- Kovacs, H., Moskau, D., and Spraul, M.: Cryogenically cooled probes - a leap in NMR technology, *Prog. Nucl. Mag. Res. Sp.*, 46, 131-155, 2005.
- Kroll, J. H., Donahue, N. M., Jimenez, J. L., Kessler, S. H., Canagaratna, M. R., Wilson, K. R., Altieri, K. E., Mazzoleni, L. R., Wozniak, A. S., Bluhm, H., Mysak, E. R., Smith, J. D., Kolb, C. E., and Worsnop, D. R.: Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol, *Nat. Chem.*, 3, 133-139, 2011.
- Kujawinski, E. B., Behn, M. D.: Automated analysis of electrospray ionization Fourier transform ion cyclotron resonance mass spectra of natural organic matter. *Anal. Chem.*, 78, 4363-4373, 2006.
- Longnecker, K., Kujawinski, E. B.: Composition of dissolved organic matter in groundwater, *Geochim. Cosmochim. Ac.*, 75, 27572-2761, 2011.
- Mao, J. D., Tremblay, L., Gagne, J. P.: Structural changes of humic acids from sinking organic matter and surface sediments investigated by advanced solid-state NMR: Insights into sources, preservation and molecularly uncharacterized components. *Geochim. Cosmochim. Ac.*, 75, 7864-7880, 2011.
- Mao, J. D., Kong, X. Q., Schmidt-Rohr, K., Pignatello, J. J., Perdue, E. M.: Advanced Solid-State NMR Characterization of Marine Dissolved Organic Matter Isolated Using the Coupled Reverse Osmosis/Electrodialysis Method. *Environ. Sci. Technol.*, 46, 5806-5814, 2012.
- Mopper, K., Stubbins, A., Ritchie, J. D., Bialk, H. M., and Hatcher, P. G.: Advanced instrumental approaches for characterization of marine dissolved organic matter: Extraction techniques, mass spectrometry, and nuclear magnetic resonance spectroscopy, *Chem. Rev.*, 107, 419-442, 2007.
- Mulder, F. A. A., and Akke, M.: Carbonyl <sup>13</sup>C transverse relaxation measurements to sample protein backbone dynamics, *Magn. Reson. Chem.*, 41, 853-865, 2003.
- Neudecker, P., Lundstrom, P., and Kay, L. E.: Relaxation dispersion NMR spectroscopy as a tool for detailed studies of protein folding, *Biophys. J.*, 96, 2045-2054, 2009.
- Pohl, C., Croot, P. L., Hennings, U., Daberkow, T., Budeus, G., and v. d. Loeff, M. R.: Synoptic transects on the distribution of trace elements (Hg, Pb, Cd, Cu, Ni, Zn, Co, Mn, Fe, and Al) in surface waters of the Northern- and Southern East Atlantic, *J. Marine Syst.*, 84, 28-41, 2011.
- Repeta, D. J., Quan, T. M., Aluwihare, L. I., and Accardi, A. M.: Chemical characterization of high molecular weight dissolved organic matter in fresh and marine waters, *Geochim. Cosmochim. Ac.*, 66, 955-962, 2002.
- Schmitt-Kopplin, P., Gabelica, Z., Gougeon, R. D., Fekete, A., Kanawati, B., Harir, M., Gebefuegi, I., Eckel, G., and Hertkorn, N.: High molecular diversity of extraterrestrial organic matter in murchison meteorite revealed 40 years after its fall, *P. Natl. Acad. Sci. USA*, 107, 2763-2768, 2010a.
- Schmitt-Kopplin, P., Gelencser, A., Dabek-Zlotorzynska, E., Kiss, G., Hertkorn, N., Harir, M., Hong, Y., and Gebefuegi, I.: Analysis of the unresolved organic fraction in atmospheric aerosols with ultrahigh-resolution mass spectrometry and nuclear magnetic resonance spectroscopy: Organosulfates as photochemical smog constituents, *Anal. Chem.*, 82, 8017-8026, 2010b.
- Simpson, A. J., Kingery, W. L., and Hatcher, P. G.: The identification of plant derived structures in humic materials using three-dimensional NMR spectroscopy, *Environ. Sci. Technol.*, 37, 337-342, 2003.
- Simpson, A. J., McNally, D. J., and Simpson, M. J.: NMR spectroscopy in environmental research: From molecular interactions to global processes, *Prog. Nucl. Mag. Res. Sp.*, 58, 97-175, 2011.



- Smernik, R. J., and Oades, J. M.: Effects of added paramagnetic ions on the  $^{13}\text{C}$  CP MAS NMR spectrum of a de-ashed soil, *Geoderma*, 89, 219-248, 1999.
- Smernik, R. J., and Oades, J. M.: Effect of paramagnetic cations on solid state  $^{13}\text{C}$  nuclear magnetic resonance spectra of natural organic materials, *Commun Soil Sci. Plan.*, 31, 3011-3026, 2000.
- Smernik, R. J., and Oades, J. M.: Paramagnetic effects on solid state  $^{13}\text{C}$  nuclear magnetic resonance spectra of soil organic matter, *J. Environ. Qual.*, 31, 414-420, 2002.
- Soule, M. C. K., Longnecker, K., Giovannoni, S. J., Kujawinski, E. B.: Impact of instrument and experiment parameters on reproducibility of ultrahigh resolution ESI FT-ICR mass spectra of natural organic matter. *Org. Geochem.*, 41, 725-733, 2010.
- Spencer, R. G.: Equivalence of the time-domain matched filter and the spectral-domain matched filter in one-dimensional NMR spectroscopy, *Concept. Magn. Reson. A.*, 36A, 255-265, 2010.
- Traficante, D. D., and Rajabzadeh, M.: Optimum window function for sensitivity enhancement of NMR signals, *Concept Magnetic Res.*, 12, 83-101, 2000.
- Tziotis, D., Hertkorn, N., Schmitt-Kopplin, P.: 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 Spectrom.*, 17, 415-421, 2011.
- Vraspir, J. M., and Butler, A.: Chemistry of marine ligands and siderophores, *Annu. Rev. Mar. Sci.*, 1, 43-63, 2009.
- Witter, A. E., Hutchins, D. A., Butler, A., and Luther, G. W.: Determination of conditional stability constants and kinetic constants for strong model Fe-binding ligands in seawater, *Mar. Chem.*, 69, 1-17, 2000.
- Ying, J. F., Grishaev, A. E., and Bax, A.: Carbon-13 chemical shift anisotropy in DNA bases from field dependence of solution NMR relaxation rates, *Magn. Reson. Chem.*, 44, 302-310, 2006.