Substantial Biogeochemical and Biomolecular Processing of Dissolved Organic Matter in an Anticyclonic Eddy in the Northern South China Sea down to Bathypelagic Depths

Miao Zhang[1,†], Siyu Li[2,†], Norbert Hertkorn*[2], Mourad Harir[2,3], Dongfeng Xu[4], Philippe Schmitt-Kopplin[2,4], Ying Wu*[1]

[1] State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai 200214, P. R. China

[2] Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Research Unit Analytical Biogeochemistry (BGC), Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany

[3] Chair of Analytical Food Chemistry, Technische Universität München, Alte Akademie 10, 85354, Freising-Weihenstephan, Germany

[4] State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, P. R. China

[†] M. Z. and S. L. have contributed equally to this work

Supplementary Material

Table S1. Sampling stations, and carbon-based bulk properties of South China Sea (SCS) seawater, which was used for PPL-based solid phase extraction (SPE) to isolate marine SPE-DOM.

Sample	Latitude	Longitude	Bottom depth (m)	DOC of seawater (μmol L^{-1})	Seawater used for SPE (L)	Extraction yield carbon (%)
L05_5m	115°36.189′ E	19°10.054′ N	2700	60.5	10	50.7
L05_80m				58.9	10	71.6
L05_200m				50.4	10	43.3
L05_400m				45.7	10	60.7
L05_1000m				40.2	10	58.8
L05_1898m				45.3	10	64.3
L11_5m	115°52.366′ E	19°50.857′ N	1514	58.6	10	49.1
L11_90m				55.7	10	36.2
L11_200m				54.2	10	40.9
L11_450m				42.6	10	52.5
L11_700m				39.8	10	53.0
L11_1490m				39.9	10	72.1

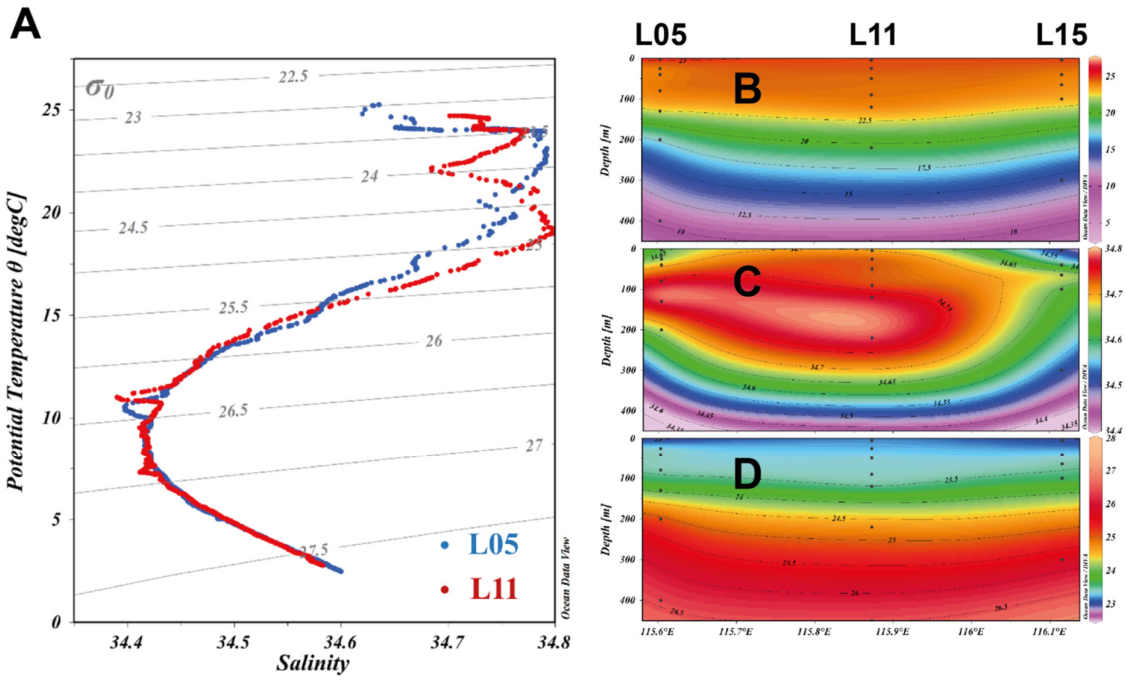

Fig. S1. (**A**) Salinity-temperature (θ-S) diagram of sampling stations, with (**B**) distribution of temperature (°C), (**C**) salinity, and (**D**) density σ_t (kg m^{-3}) across sampling path section X (cf. Fig. 1).

Sampling and analysis of fluorescent and chromophoric DOM

Fluorescent and chromophoric dissolved organic matter (FDOM and CDOM) samples were filtered using PES syringe filters (Millipore, 0.22 μm pore size) into 20 mL glass ampoules, precombusted in 450 ℃ for 5 h.

Fluorescence spectra were determined by a fluorescence spectrophotometer (HITACHI F-4500) with a 1 cm quartz cuvette. Emission (Em) spectra were scanned from wavelength 300 nm to 550 nm with an interval of 1 nm, while excitation (Ex) spectra were recorded every 6 nm from 274 nm to 478 nm. Slit widths of both Ex and Em were 5 nm. The excitation emission matrix spectra (EEMs) were corrected by subtracting the Milli-Q water EEMs regarded as blank, which were scanned every day. EEMs were Raman normalized also with Milli-Q water EEMs scanned every time (Zepp et al. 2004). Inner filter effects were ignored because of the low absorbance of open ocean water samples (Kothawala et al. 2013). The components were identified from EEMs using parallel factor analysis (PARAFAC) model with MATLAB 2014 as well as the DOMFluor toolbox (Murphy et al. 2013, Stedmon and Bro 2008). The freshness index (FrI) was computed as the ratio of emission intensity at 380 nm to the maximum emission intensity between 420 and 435 nm at the excitation wavelength 310 nm (Parlanti et al. 2000).

CDOM absorption spectra were determined with an ultraviolet-visible spectrophotometer (Pgeneral TU-1901). 10 cm quartz cuvettes were used because of the low concentration of seawater. Blank was calibrated with Milli-Q water once every 5 samples. Absorption spectra were recorded from wavelengths 800 nm to 200 nm with an interval of 1 nm. Baseline was corrected by subtracting the average absorption between 700 nm and 710 nm. CDOM absorption coefficients (a(λ)) were computed using equation (1):

$$a(\lambda) = 2.303 \times A(\lambda)/L \quad (1)$$

where A(λ) is the absorbance at wavelength λ nm. L was the optical path length (10 cm). a_{325} was the absorption coefficient at wavelength 325 nm, representing the concentration of CDOM in the open ocean (Nelson and Gauglitz 2016, Nelson et al. 2010). Specific ultraviolet absorbance (SUVA) is defined as A(λ) normalized for DOC concentration (C_{DOC}). $SUVA_{254}$ (L mg^{-1} m^{-1}) represented the UV absorbance at 254 nm divided by the C_{DOC}, which is indicative of the aromaticity of DOM (Weishaar et al. 2003).

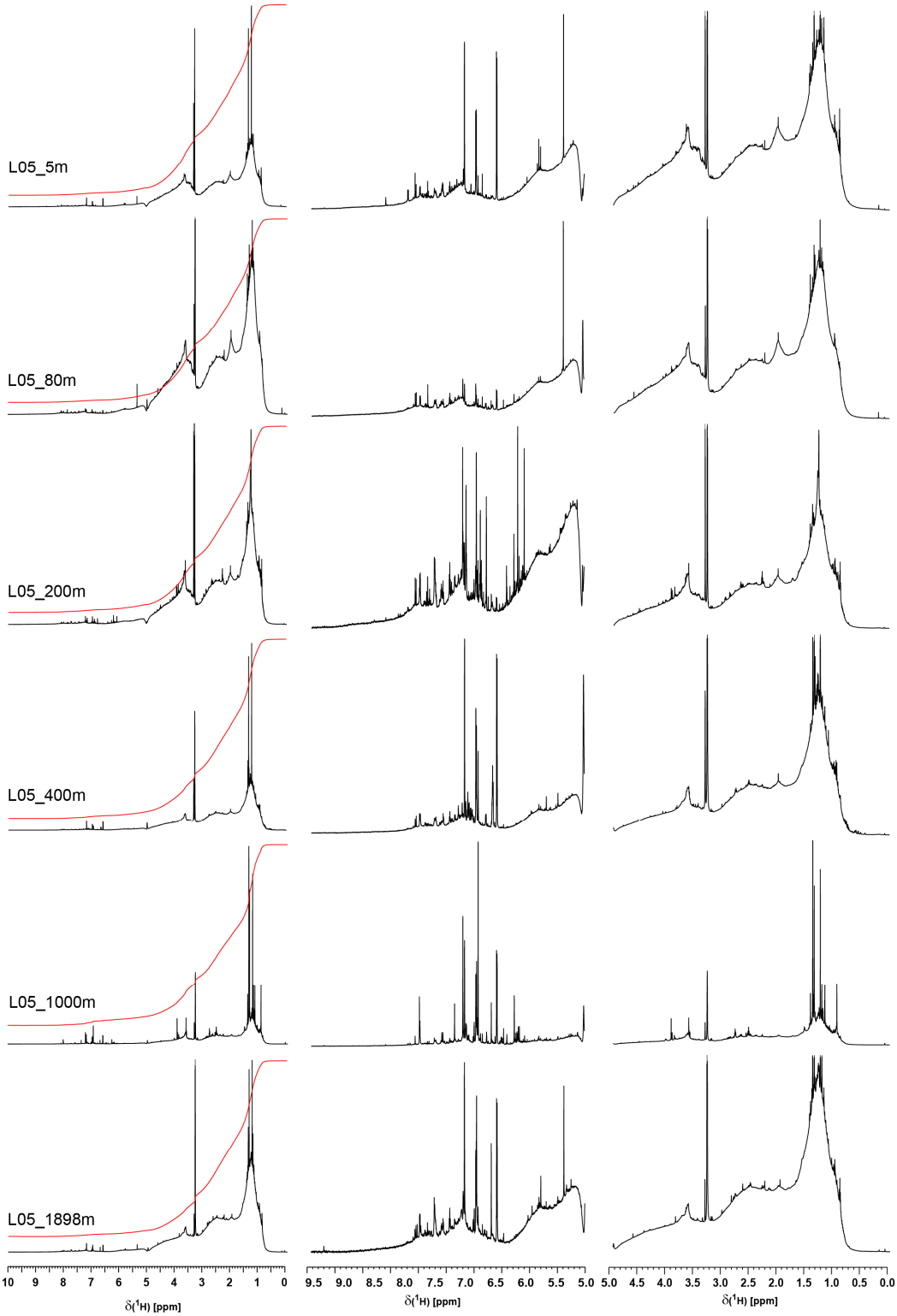

Fig. S2. ^{1}H NMR spectra (800 MHz, CD_3OD) of South China Sea SPE-DOM (PPL): depth profiles for station L05.

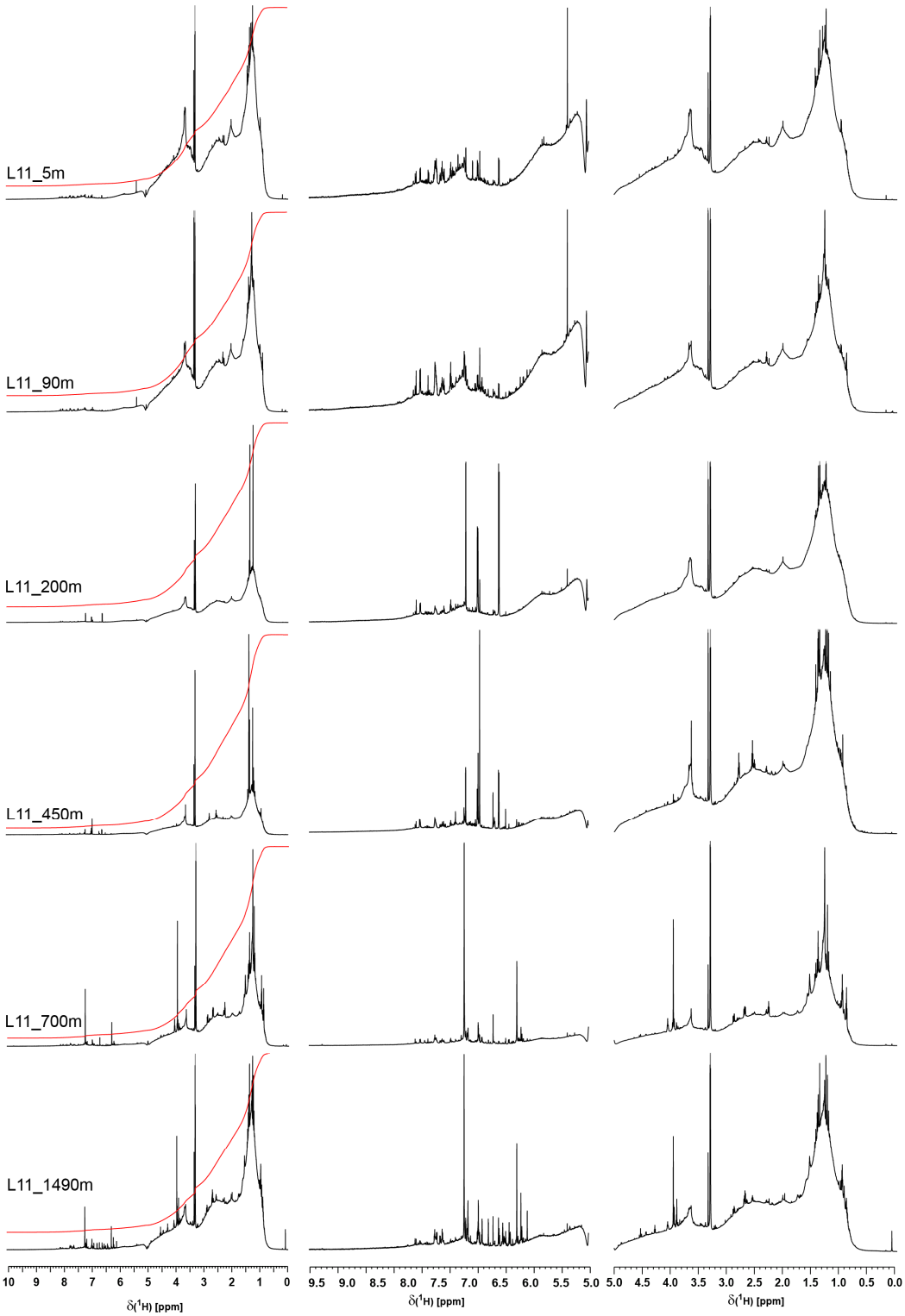

Fig. S3. ^{1}H NMR spectra (800 MHz, CD_3OD) of South China Sea SPE-DOM (PPL): depth profiles for station L11.

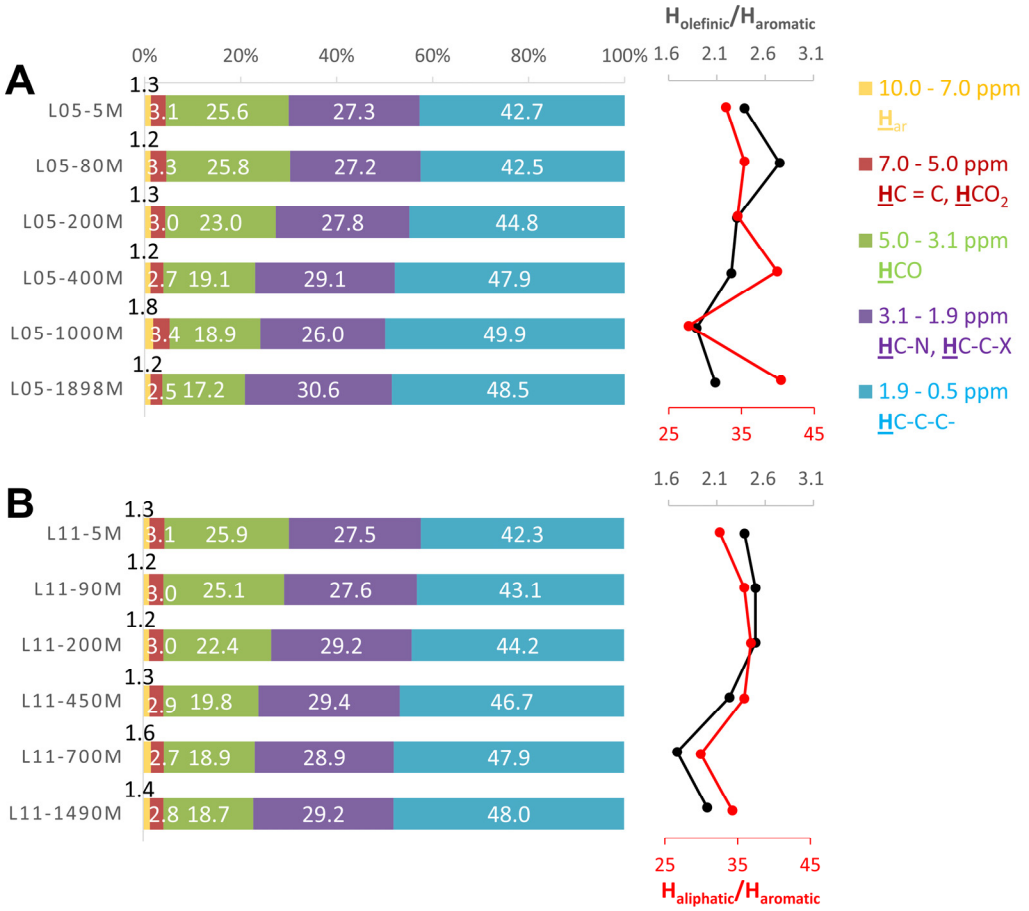

Fig. S4. [1]H NMR section integral for key substructures of sampling stations (**A**) L05 and (**B**) L11; the black line indicates the ratio between olefinic and aromatic compounds; the red line refers to the ratio between aliphatic and aromatic compounds.

Table S2. Acquisition parameters of NMR spectra, shown according to figures. PK: probeheads used for acquisition of NMR spectra, 8QCI: cryogenic inverse geometry 5 mm z-gradient $^{1}H/^{13}C/^{15}N/^{31}P$ QCI probe (B_0 = 18.8 T); 5D: cryogenic classical geometry 5 mm z-gradient ^{13}C, ^{1}H probe (B_0 = 11.7 T); 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 is given in [Hz], SI/QS derived functions indicate shift by π/n. (*) asterisk: coadded ^{13}C NMR spectra.

spectrum	Figure	PK	NS	AQ [ms]	D1 [ms]	NE	WDW1	WDW2	PR1	PR2
^{1}H NMR	2, S2, S3	8QCI	224-1280	5000	5000	-	EM	-	1	-
^{13}C DEPT-135 NMR	4, S15, S16	5D	82944*	1000	2000	-	EM	-	12.5	-
^{13}C NMR	4, S15, S16	5D	83456*	1000	19000		EM	-	12.5	
^{1}H, ^{1}H JRES	S15A, S16A	8QCI	80	1000	500	64	QS	QS	0	0
^{1}H, ^{1}H COSY	S18CD	8QCI	64	750	2250	1600	EM	QS	2.5	2.5
^{1}H, ^{1}H TOCSY	S15B, S16B	8QCI	4	1000	2500	1600	EM	QS	2.5	4
^{1}H, ^{1}H NOESY	S16C	8QCI	8	1000	2500	1600	EM	QS	2.5	4
^{1}H,^{13}C DEPT HSQC	4, 6AB, S15C, S16D	8QCI	400	250	1250	401	EM	QS	7.5	2.5
^{1}H, ^{13}C HMBC	4C, S15C, S16E	8QCI	640	1000	500	162	EM	QS	2.5	2.5

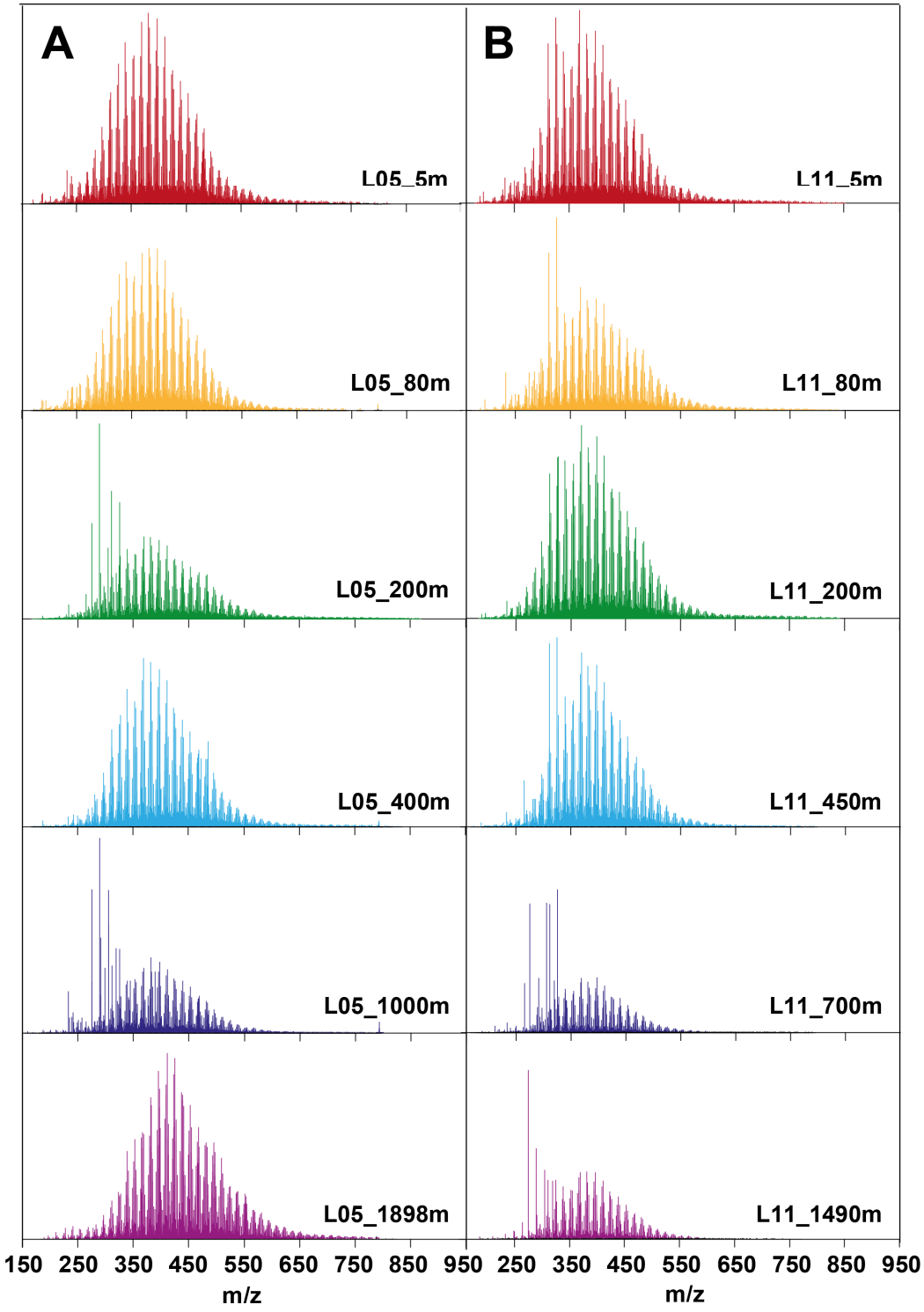

Fig. S5. Negative electrospray ionization FTICR mass spectra of South China Sea SPE-DOM; (**A**) from 5-1898 m in station L05 and (**B**) from 5-1490 m in station L11. Spectra show the molecular ions of the carboxylic acid derivative **M2** ($C_{16}H_{21}O_4^-$; m/z 205.15924) which ionized better than the phenol **M1** ($C_{14}H_{21}O^-$; m/z 277.14398) as expected. The relative abundance of **M1**/**M2** in ^{1}H NMR spectra and in FTICR mass spectra grossly followed each other, but contributions of many isomers with same molecular formula, and ionization selectivity will interfere with exact congruence of these complementary data.

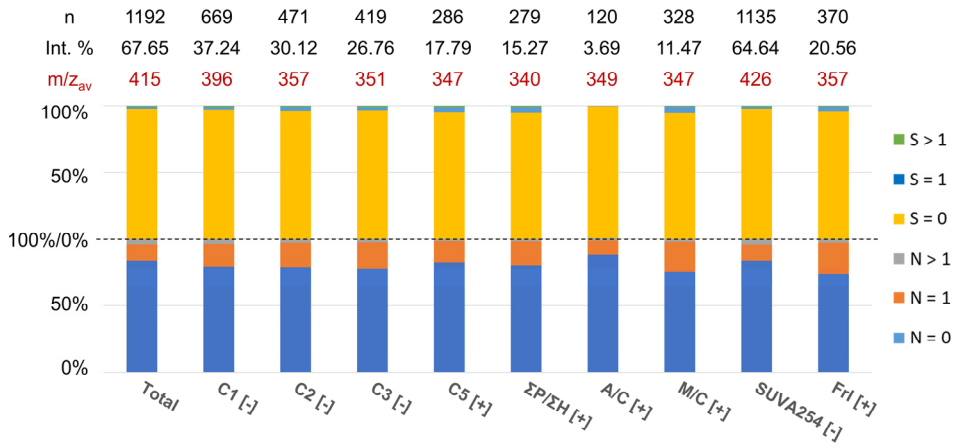

Fig. S6. Color-coded counts of nitrogen and sulfur atoms within selected, FTMS-derived classes of compounds which were correlated positively [+] or negatively [-] ($p < 0.05$) with optical parameters, referred to normalized mass spectra. n: count of molecular compositions, which related significantly with optical parameters; Int. %: relative percentage of mass peak amplitude according to class of compounds; m/z (Da): average molecular weight of each class of compounds.

Table S3. FTMS-derived categorizations of selected assigned classes of compounds in SCS SPE-DOM (cf. Fig. S6), with counts of assigned compounds, and FTMS-derived relative mass peak intensities per class.

	L05 5m	**L05 80m**	**L05 200m**	**L05 400m**	**L05 1000m**	**L05 1898m**	**L11 5m**	**L11 90m**	**L11 200m**	**L11 450m**	**L11 700m**	**L11 1490m**
Condensed compounds	163	40	124	129	119	138	112	140	113	116	156	75
	1.34%	0.34%	0.74%	1.05%	0.75%	1.05%	0.79%	0.89%	0.58%	0.80%	1.49%	0.66%
Highly unsaturated compounds	2350	1921	2896	2484	2505	3061	2610	3046	2670	2282	2087	1992
	84.9%	86.6%	86.0%	87.7%	83.3%	91.4%	85.6%	84.4%	87.4%	87.3%	82.4%	88.0%
Unsaturated compounds	745	559	756	634	652	498	771	936	662	501	580	459
	10.60%	10.47%	9.79%	8.23%	10.89%	5.16%	10.70%	11.67%	8.91%	8.25%	10.89%	7.32%
Peptides-like	297	227	337	257	652	183	333	444	244	187	201	149
	2.84%	2.88%	2.58%	2.29%	3.47%	1.12%	2.81%	3.18%	1.94%	1.78%	1.74%	1.32%
Saturated compounds	32	13	27	31	16	15	18	46	17	18	36	24
	0.51%	0.14%	0.51%	0.48%	0.27%	0.13%	0.24%	0.59%	0.16%	0.51%	1.26%	0.56%
Sugars-like	42	26	40	32	34	16	34	49	36	24	16	21
	0.32%	0.23%	0.22%	0.22%	0.22%	0.11%	0.19%	0.21%	0.21%	0.14%	0.09%	0.16%

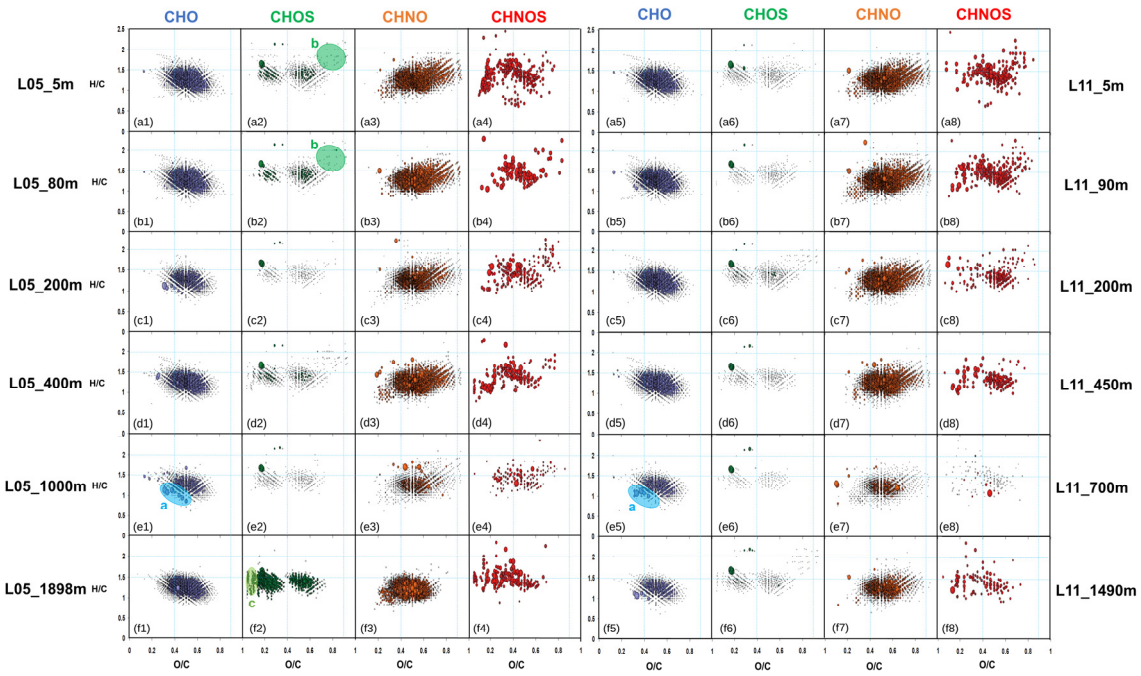

Fig. S7. H/C versus O/C van Krevelen diagrams of SPE-DOM in L05 and L11, obtained from negative electrospray 12T FTICR mass spectra. From top to bottom: L05_5 m (a1-a4), L11_5 m (a5-a8), L05_80 m (b1-b4), L11_90 m (b5-b8), L05_200 m (c1-c4), L11_200 m (c5-c8), L05_400 m (d1-d4), L11_450 m (d5-d8), L05_1000 m (e1-e4), L11_700 m (e5-e8), L05_1898 m (f1-f4), L11_1490 m (f5-f8). Color code is according to molecular series as follows: CHO, blue; CHOS, green; CHNO, orange; CHNOS, red. Bubble areas reflect the relative mass peak intensity. Section *a*: Less oxygenated more saturated CRAM; section *b*: S-bearing carbohydrates molecular series; section *c*: Less oxygenated CHOS molecular series.

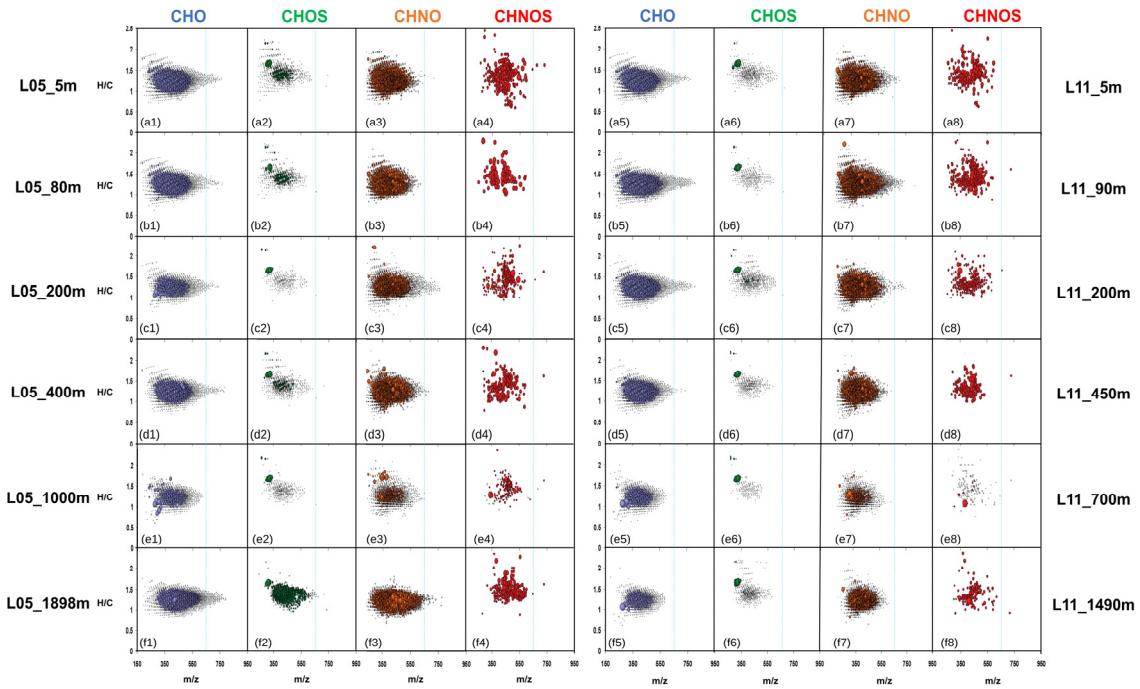

Fig. S8. Mass edited H/C ratios derived from negative ESI FTICR mass spectra of SPE-DOM in L05 and L11. From top to bottom: L05_5 m (a1-a4), L11_5 m (a5-a8), L05_80 m (b1-b4), L11_90 m (b5-b8), L05_200 m (c1-c4), L11_200 m (c5-c8), L05_400 m (d1-d4), L11_450 m (d5-d8), L05_1000 m (e1-e4), L11_700 m (e5-e8), L05_1898 m (f1-f4), L11_1490m (f5-f8). Color coded according to molecular series as follows: CHO, blue; CHOS, green; CHNO, orange; CHNOS, red.

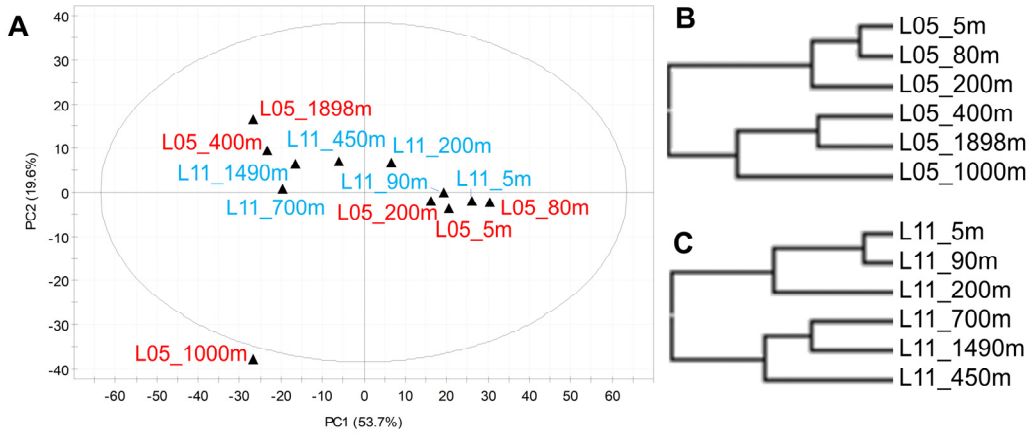

Fig. S9. NMR-based (800 MHz ^{1}H NMR, CD_3OD, area-normalized from 0.5-10.0 ppm; 0.01 ppm bucket resolution) (**A**) principal component analysis (PCA) and (**B**) hierarchical cluster analysis (HCA) of stations L05 and L11 depth profiles.

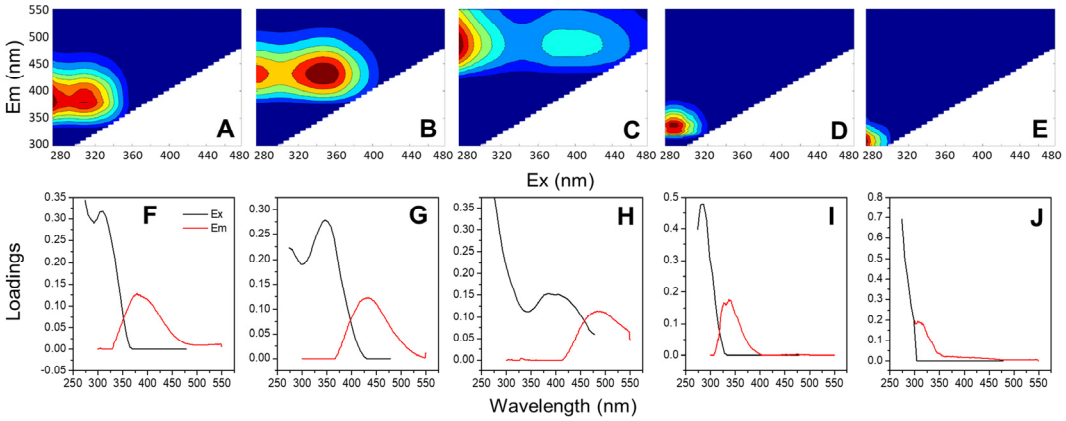

Fig. S10. EEM contours and excitation (black line) and emission (red line) of five PARAFAC components. (**A**, **F**) C1 (peak M) with Ex/Em maxima of <270, (310)/380 nm, (**B**, **G**) C2 (Ex/Em maxima = <270, (350)/435 nm), with similar characteristics with fluorophore C, (**C**, **H**) C3 with Ex/Em maxima of <250, (400)/490 nm, (**D**, **I**) C4 with Ex/Em maxima of <280, 340 nm and (**E**, **J**) C5 with Ex/Em maxima of <280/310 nm components.

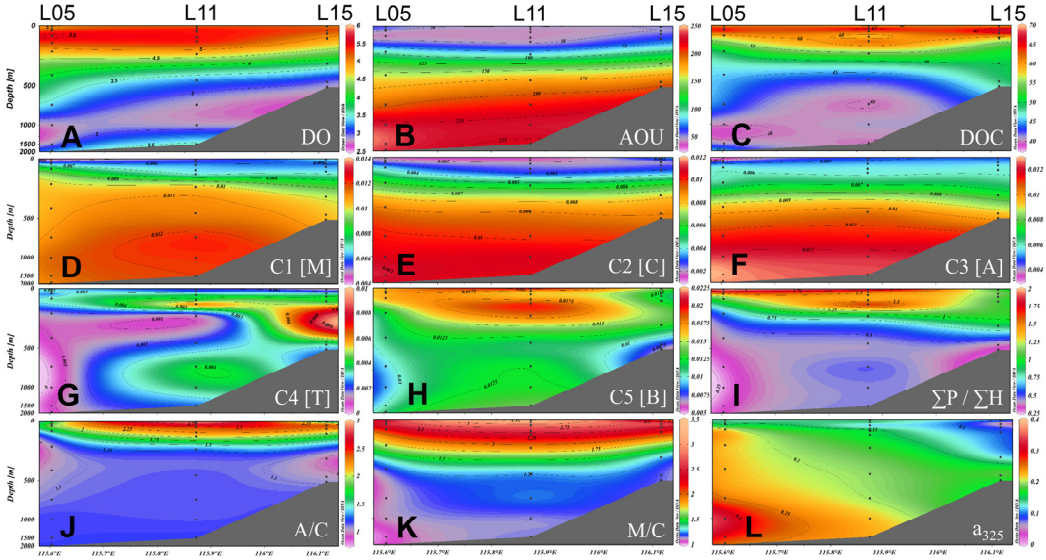

Fig. S11. The distribution of (**A**) DO (mg L^{-1}), (**B**) AOU (μmol L^{-1}), (**C**) DOC (μmol L^{-1}), (**D**) C1 (R.U.), (**E**) C2 (R.U.), (**F**) C3 (R.U.), (**G**) C4 (R.U.), (**H**) C5 (R.U.), (**I**) ΣP/ΣH (**J**) A/C, (**K**) M/C (**L**) a_{325} (m^{-1}) (**L**) in section X.

Table S4. The correlation coefficient (R) between AOU (μmol L^{-1}), DOC (μmol L^{-1}) and fluorescent components (C1-C5, R.U.), a_{325} (m^{-1}) of sampling stations L05 and L11.

Stations	R	C1	C2	C3	C4	C5	a_{325}
L05	AOU	0.86**	0.97**	0.98**	-0.85**	-0.92**	0.84**
	DOC	-0.94**	-0.97**	-0.93**	0.82**	0.96**	-0.91**
L11	AOU	0.93**	0.99**	0.99**	-0.10	-0.80**	0.82**
	DOC	-0.93**	-0.96**	-0.95**	0.17	0.74*	-0.80**

** correlation is significant at the 0.01 level.

* correlation is significant at the 0.05 level.

Fluorescent and chromophoric dissolved organic matter properties

Five components were separated from excitation-emission matrix spectra (EEMs) using PARAFAC analysis, including three humic-like components (C1, C2, C3) and two protein-like components (C4, C5) (Fig. S10). All five components were matched to the previous results by testing with OpenFluor database (Murphy et al. 2014b). C1 (peak M) (Coble 1996) with Ex/Em maxima of <270, (310)/380 nm (Figs. S11a, f), was considered as microbial-derived substances or byproducts of photo-bleaching of terrestrial materials (Gonçalves-Araujo et al. 2016, Li et al. 2015, Osburn et al. 2016, Tanaka et al. 2014). C2 (Ex/Em maxima = <270, (350)/435 nm), with similar characteristics with fluorophore peak C (Figs. S10b, g) (Coble 1996), were generally regarded as terrestrially derived substance (Guéguen et al. 2014, Osburn et al. 2015, Walker et al. 2009). C3 (Peak A) (Coble 1996) was categorized as high molecular weight and aromatic organic compounds, which was widespread in aquatic environment, with maxima Ex/Em maxima = <250, (400)/490 nm (Figs. S10c, h) (Amaral et al. 2016, Catalá et al. 2015, Cawley et al. 2012, Murphy et al. 2014a, Osburn et al. 2016). Humic-like components were more resistant to degradation than protein-like components, which were accumulated from surface to deep (Figs. S10d, e, f). C4 and C5 components (peaks T and B) (Coble 1996) were considered as tryptophan- and tyrosine-like components with higher bioavailability (Osburn and Stedmon 2011, Painter et al. 2018, Podgorski et al. 2018, Yamashita et al. 2013, Yamashita et al. 2011). C4 and C5 components could be produced by photosynthesis of phytoplankton in the euphotic zone and be degraded progressively, and hence decreased from surface to bottom (Figs. S10g, h).

The ratio of total protein and total humic components [ΣP/ΣH = (C4+C5)/(C1+C2+C3)] indicated the contribution of bioavailable organic matter to total DOM (Coble 2007), which decreased with depth (Fig. S8i). A/C was defined as ratio of the intensity of peaks A (C3) and C (C2) while M/C was ratio between peaks M (C1) and C (C2). Higher A/C represented more DOM was photodegraded (Coble 1996), while higher M/C indicated more marine source of total DOM (Helms et al. 2013). In sampling cross section X (Fig. 1), the impact of downwelled anticyclonic eddy was obvious, especially on DO, AOU, DOC, C1, C2, C5, ΣP/ΣH, A/C and M/C in the upper 400 m, whose isolines were parallel with temperature, salinity and density (Figs. S1, S8). Not all the organic substances, such as a_{325}, C3 and C4, distributed consistently

with the downwelling, which could be caused by multiple effects of biochemical and physical processes (Fig. S8).

The FDOM components showed strong correlation with AOU and DOC in L05 and L11 (Table. S4). Humic-like components C1, C2, C3 related positively with AOU and negatively with DOC, which further suggested links between remineralization and formation of humic substances (Nelson and Gauglitz 2016). The protein materials of L05 showed strong negative correlation with AOU and positive correlation with DOC, which could be a consequence of microbial utilization of a less recalcitrant fraction which in amino acid-like substances (Catalá et al. 2016). However, the relationship between protein components and AOU were not significant in L11 station, especially tryptophan-like composition C4.

Because of photo bleaching, a_{325} was lower on surface and higher in the deep, while in L05 was higher than L11 (Fig. S11l). Strong negative linear correlation was observed between DOC and a_{325} with R = -0.91 ($p < 0.01$) for L05 and R = -0.80 ($p < 0.01$) for L11. Also, a_{325} was correlated positively with AOU (L05: R = 0.84, L11: R = 0.82) (Table S4), indicating more than 80% CDOM could be related to the organic matter oxidation, which was similar as found in Pacific and Indian Ocean (Nelson et al. 2010).

Linking optical and molecular properties of SCS DOM

Spearman's rank correlations (95% confidence limit) between normalized (Max-min normalization) mass spectral peak intensities and fluorescence intensities, fluorescence indices and absorption coefficients in SCS were carried out to elucidate groups of molecules associated with optical properties. These results indicated that about 1200 assigned compounds with >67% of total mass peak intensities were correlated positively (R > 0.58) or negatively (R < -0.58) with one or more optical parameters (Table S5). This percentage was larger than analogous results found in Canadian boreal rivers (59%) (Stubbins et al. 2014), but lower than those in Florida Everglades (82%) and the Mediterranean Sea (70%) (Martínez-Pérez et al. 2017, Wagner et al. 2015). The most significant molecular series correlating with optical parameters were highly unsaturated (64%) and unsaturated compounds (3%), whereas barely condensed, saturated and sugars-like compounds relating with chromophoric and fluorescent DOM (Figs. S12, S13). Molecules correlating with optical parameters of DOM were mostly CHO compounds (55.2% of the total intensity), ~400 CHNO compounds (~10% of intensity), with only a few CHOS compounds contributing (~1.6% of intensity) (Fig. S12; Table S5). HU-type CHNO molecules correlated with optical parameters of DOM, whereas S-containing series were HU and UC with less oxygenation (Fig. S12).

Humic-like fluorescent components C1, C2, C3 (discussed in supplementary section) correlated negatively with CHO, CHOS, CHNO and CHNOS molecular series. Others such as protein-like fluorescent component C5, ΣP/ΣH, M/C, FrI (discussed in supplementary), which could indicate the freshness and (semi-) labile DOM showed positive correlation with assigned molecular series. Only 4 formulas (0.1% of intensity) related significantly with tryptophan-like component C4 (peak T)(Table S5), which could be caused by the selective retention of PPL cartridges and ionization selectivity in (-)ESI (Reemtsma et al. 2008). Approximately 500 molecules representing more than 29% of spectral intensity correlated negatively with the aromaticity indicator $SUVA_{254}$ (Table S5). Smaller molecules were more likely to associate with optical parameters (Fig. S13), most of which were less than 370 Da. More molecular series correlated with the optical variables that associated with humification and refractory character of DOM (C1, C2, C3, $SUVA_{254}$), with higher molecular weight (Fig. S13). The results were coincident with that found in Mediterranean Sea (Martínez-Pérez et al. 2017). The ratio between peak

A (C3) and peak C (C2) could indicate the photo-degradation, which correlated with molecular series without N or S (Table S5) and low molecular weight (Fig. S13), indicating photo-bleaching had less impact on dissolved organic nitrogen and sulfur, and had strong effect on small molecules. The molecular series correlated with ΣP/ΣH, M/C, FrI contained more N and S with low molecular weight (Fig. S3; Table S5), which represented heteroatom-rich as well as small molecules in newly-produced DOM.

The Spearman's rank correlation between ^{1}H NMR spectra and optical parameters provided that about 86.9% of NMR resonances related with CDOM and FDOM (Table S6). 87.3% of saturated protons and 76.8% of unsaturated protons showed significant correlations with optical properties. Most protons related with one or more optical parameters (Table S6; Fig. S14). More than three quarters of protons had strong correlation with the humic-like fluorescent components (C1, C2, C3) while 65.6% of protons related with protein-like compounds (C5). Similar to molecular series of FTICR/MS, less protons (26.8%) showed significant correlation with C4. Compared with saturated protons, more unsaturated protons related with more labile compounds, such as C5, ΣP/ΣH, M/C (Table S6). The intensity ratio of unsaturated protons that correlated with A/C was higher than that of saturated protons, indicating the unsaturated protons were more photodegradable (Table S6). Aliphatic protons were correlated with humic-like FDOM and $SUVA_{254}$ positively, while singly oxygenated protons and unsaturated protons related with low-bioavailable compounds negatively (Figs. S14a, b, c, h). The correlations between different protons with the parameters of fresher DOM were inverse. The positive correlation of unsaturated protons and bioavailable compounds demonstrated the biological source of olefins (Figs. S14d, e, g, i). The **H**CO protons such as carbohydrates, ethers, esters and olefins were more likely to be photodegraded in the ocean (Fig. S14f).

Table S5. The number of assigned molecular formulas with 0, 1 and more nitrogen or sulfur atoms which correlated positively [+] and negatively [-] with optical parameters. The numbers in parentheses represented the percentage of intensity from mass spectra.

	Total	**N = 0**	**N = 1**	**N > 1**	**S = 0**	**S = 1**	**S > 1**
Total correlated molecular series	1192 (67.65%)	787 (56.81%)	261 (8.21%)	144 (2.63%)	1117 (66.06%)	58 (1.15%)	17 (0.45%)
C1 [-]	669 (37.24%)	422 (29.66%)	183 (6.25%)	64 (1.33%)	608 (35.99%)	49 (0.95%)	12 (0.29%)
C2 [-]	471 (30.12%)	277 (23.94%)	154 (5.43%)	40 (0.82%)	414 (28.93%)	48 (1.01%)	9 (0.24%)
C3 [-]	419 (26.76%)	238 (20.95%)	147 (5.12%)	65 (0.69%)	372 (25.75%)	39 (0.78%)	8 (0.23%)
C4 [+]	4 (0.10%)	4 (0.10%)	0 (0)	0 (0)	4 (0.10%)	0 (0)	0 (0)
C5 [+]	286 (17.79%)	186 (14.72%)	82 (2.79%)	18 (0.28%)	243 (18.12%)	37 (0.74%)	6 (0.16%)
ΣP/ΣH [+]	279 (15.27%)	176 (12.31%)	85 (2.70%)	18 (0.27%)	241 (14.45%)	34 (0.68%)	5 (0.14%)
A/C [+]	120 (3.69%)	93 (3.14%)	21 (0.38%)	6 (0.06%)	115 (3.53%)	5 (0.01%)	0 (0)
M/C [+]	328 (17.18%)	191 (12.95%)	115 (3.87%)	22 (0.36%)	286 (16.23%)	38 (0.84%)	4 (0.11%)
$SUVA_{254}$ [-]	1135 (64.64%)	743 (54.24%)	248 (7.79%)	143 (2.61%)	1064 (63.16%)	55 (1.04%)	16 (0.44%)
FrI [+]	370 (20.56%)	205 (15.17%)	137 (4.82%)	28 (0.57%)	327 (19.65%)	39 (0.81%)	4 (0.10%)

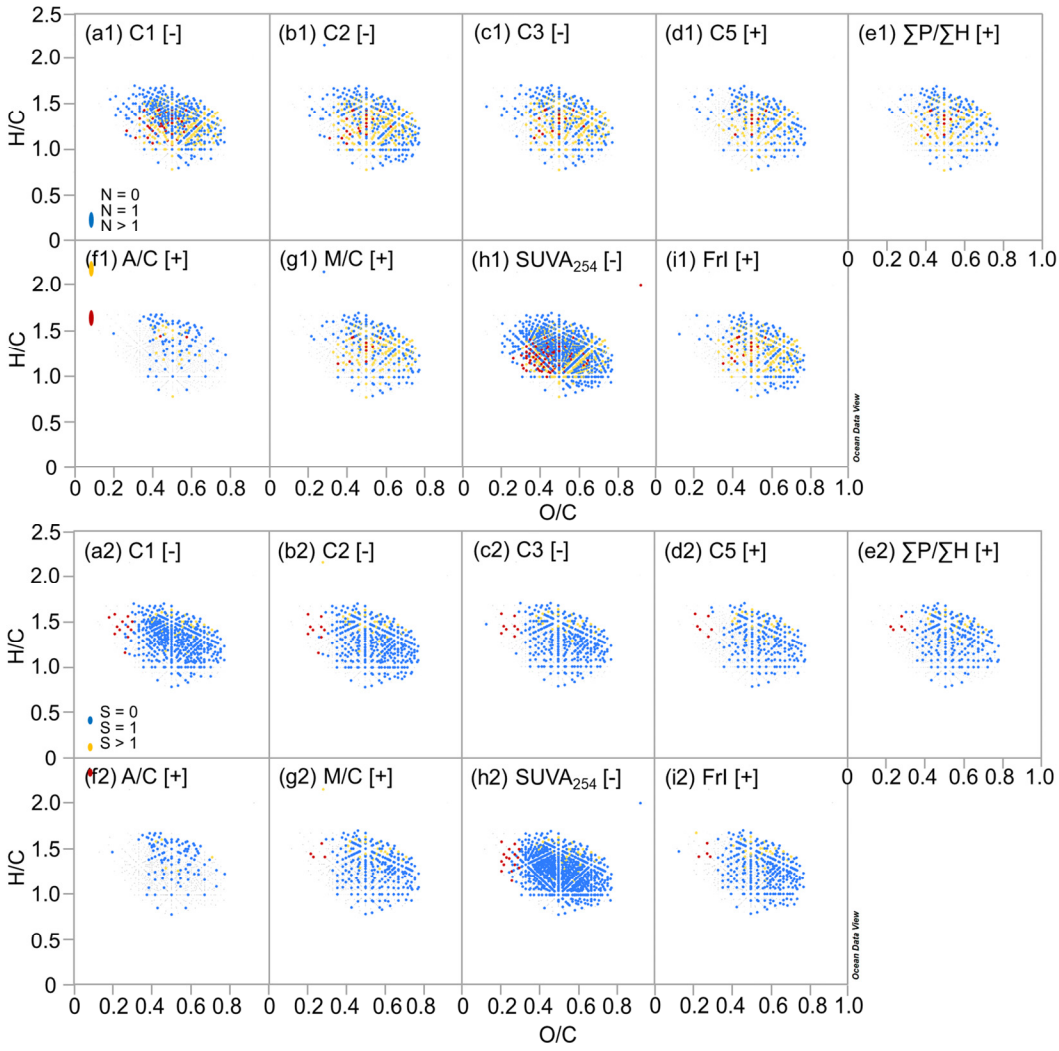

Fig. S12. H/C versus O/C van Krevelen diagrams of SPE-DOM in SCS SPE-DOM. Molecular series correlated with FDOM components C1 (R.U.) (a), C2 (R.U.) (b), C3 (R.U.) (c), C5 (R.U.) (d), ΣP/ΣH (e), A/C (f), M/C (g), $SUVA_{254}$ (L mg^{-1} m^{-1}) (h), FrI (i) in 95% confidence limit. The color of the dots indicated the number of nitrogen (a1-i1) or sulfur (a2-i2). The plus-minus sign [+] [-] represented the correlations were positive and negative, respectively.

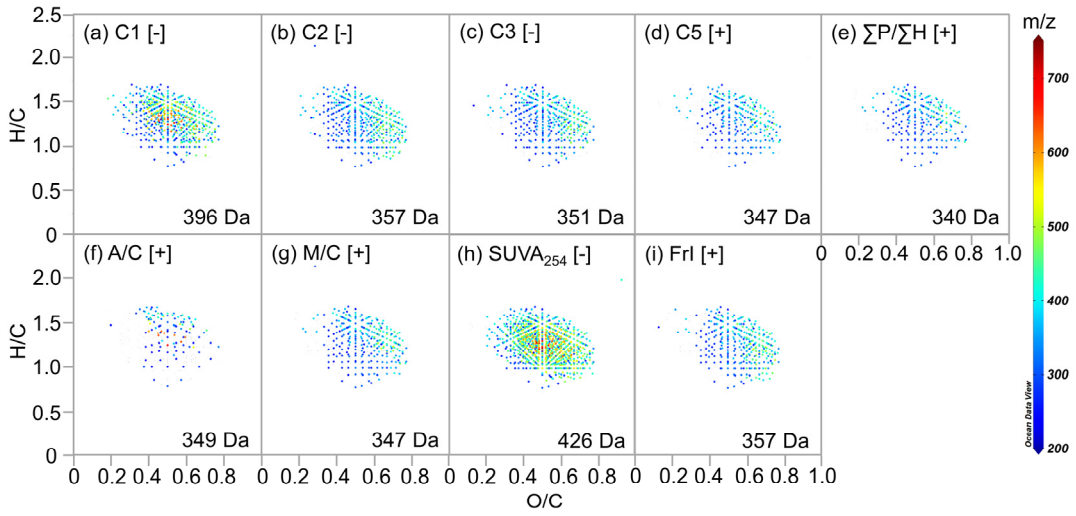

Fig. S13. H/C versus O/C van Krevelen diagrams of SPE-DOM in SCS. Molecular series correlated with FDOM components (a) C1 (R.U.), (b) C2 (R.U.), (c) C3 (R.U.), (d) C5 (R.U.), (e) ΣP/ΣH, (f) A/C, (g) M/C, (h) $SUVA_{254}$ (L mg^{-1} m^{-1}), (i) FrI in 95% confidence limit. The plus-minus signs [+] [-] represented the respective positive and negative correlations. The average molecular weight is provided in the right-bottom corner of each panel. Color indicated the respective weight of molecular series.

Table S6. The signal number of ^{1}H NMR spectra (0.01 ppm bucket resolution) correlated with optical parameters. The numbers in parentheses represented the intensity percentage of each type of proton (e.g. all protons, saturated protons, unsaturated protons).

	Total	Saturated proton δ_H < 5.0 ppm	Unsaturated proton δ_H > 5.0 ppm
Total correlated NMR signal	70 (86.9%)	34 (87.3%)	37 (76.8%)
C1	60 (81.2%)	31 (81.8%)	29 (67.4%)
C2	61 (75.7%)	29 (75.9%)	32 (71.2%)
C3	62 (79.8%)	31 (80.3%)	31 (69.5%)
C5	56 (65.6%)	25 (65.4%)	31 (70.0%)
ΣP/ΣH	63 (71.9%)	28 (71.8%)	35 (73.4%)
A/C	27 (43.8%)	17 (43.7%)	10 (46.2%)
M/C	63 (71.8%)	28 (71.8%)	35 (72.5%)
$SUVA_{254}$	43 (63.6%)	24 (63.6%)	19 (63.6%)
FrI	56 (73.8%)	28 (74.2%)	28 (63.8%)

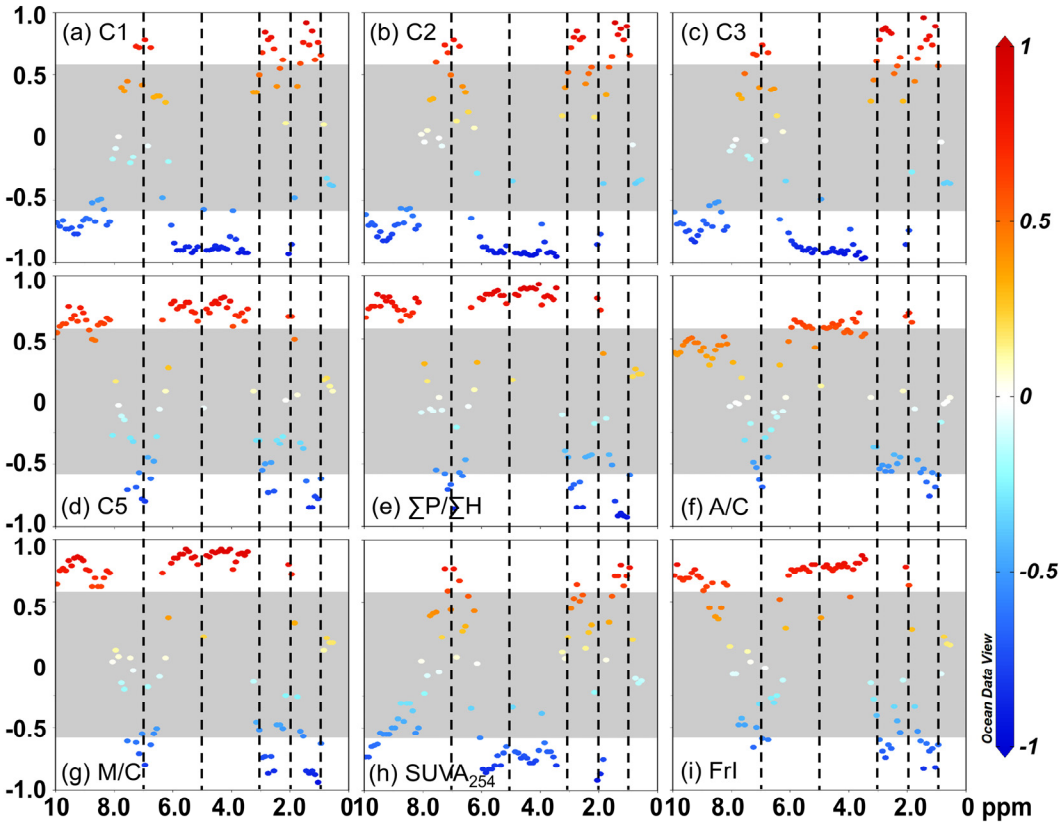

Fig. S14. The R value (Spearman's rank correlation) of ^{1}H NMR spectra signal (0.01 ppm bucket resolution) and optical parameters. The dots within the gray shadow represented $p > 0.05$, indicating the correlation was not significant.

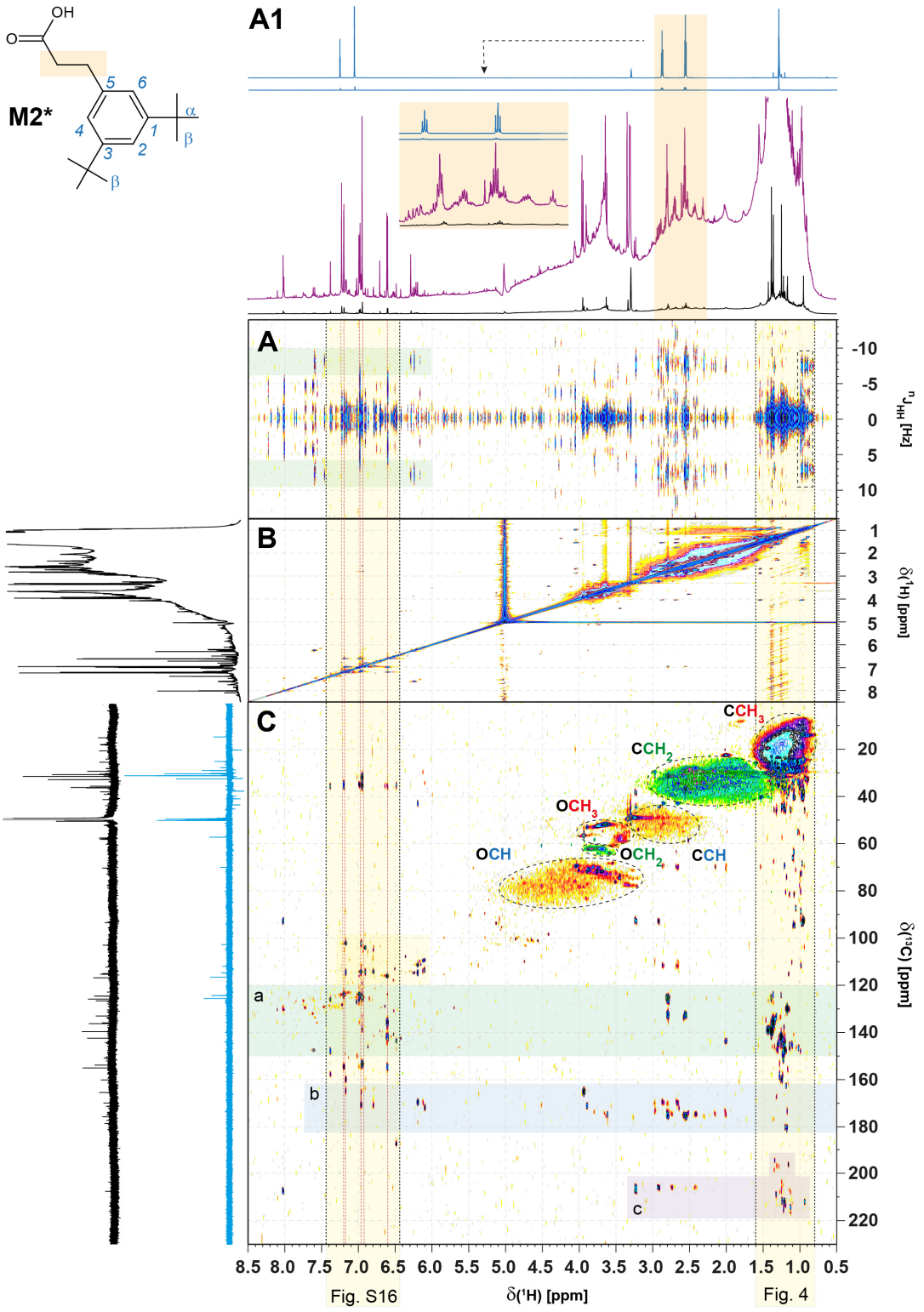

Fig. S15. Full caption see next page.

Fig. S15. ^{1}H NMR detected NMR spectra of L05_1000 m SPE-DOM, showing characteristics of common biogeochemical marine organic matter and those of rather abundant small molecules. The ^{1}H NMR chemical shift sections depicted in Fig. 4 and Fig. S17 are highlighted in yellow. (**A1**): ^{1}H NMR spectrum of compound M2* (cf. Fig. S16), with shaded insert denoting HOOC-C**H**$_2$-C**H**$_2$-groups; panel (**A**): ^{1}H, ^{1}H JRES; (**B**): ^{1}H, ^{1}H TOCSY; panel (**C**): overlay of ^{1}H, ^{13}C DEPT HSQC NMR and ^{1}H, ^{13}C HMBC NMR spectra, with ^{13}C-DEPT-135 (blue) and single pulse ^{13}C NMR spectra (black); 2D NMR spectra show an abundant and widely distributed assortment of overall ~400 intense sharp cross peaks, probably representing ~ 75 rather abundant biological (metabolites) or chemical (e.g. pollutants) molecules, with tentative assignment provided (cf. text). ^{1}H, ^{13}C HMBC NMR cross peaks with gross counts (z) and tentative assignment. Section a (and b): polarized trans-double bonds (=C-**H**C=C**H**; z ~ 7); section b: **C**$_{ar}$C$_n$**H** (n ≥1; z ~ 200); section c: HOO**C**C$_n$**H** (n ≥1; z ~ 65); section d: -C-C(=O)C$_n$**H** (n ≥1; z ~ 50). ^{13}C NMR spectra at Fig. S15C are ^{13}C single pulse (black) and CH_2-selective (blue) ^{13}C DEPT NMR spectra.

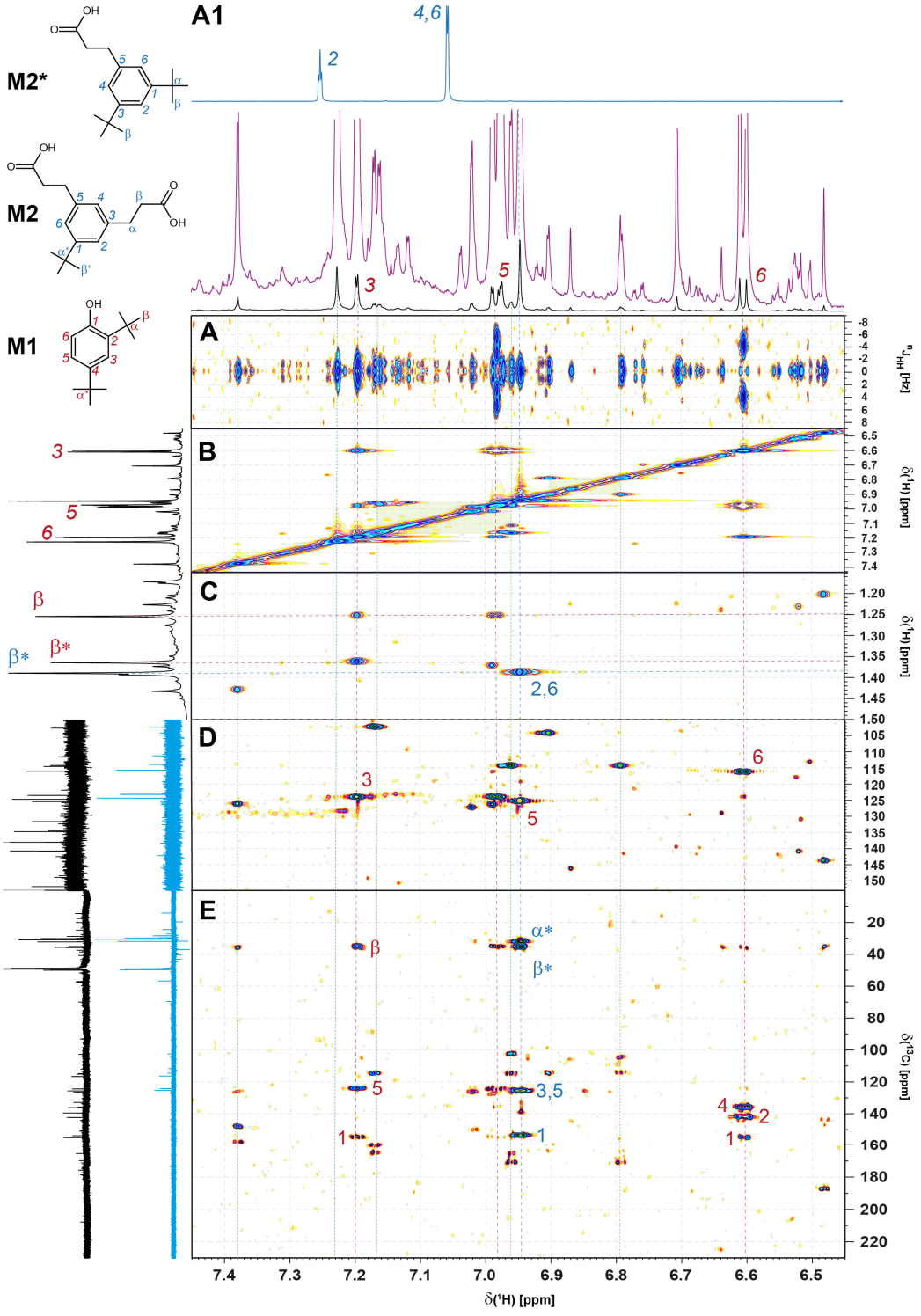

Fig. S16. ^{1}H NMR detected NMR spectra of $C_{ar}H$ groups of L05_1000 m SPE-DOM; (**A**): ^{1}H, ^{1}H JRES; panel (**B**): ^{1}H, ^{1}H TOCSY; panel (**C**): ^{1}H, ^{1}H NOESY NMR spectrum, with cross peaks denoting spatial proximity between $C_{ar}H$ and tert-butyl groups C_4H_9; panel (**D**): ^{1}H, ^{13}C DEPT-HSQC (red: CH_3; green CH_2) and (panel E) ^{1}H, ^{13}C HMBC NMR spectra of L05_1000 m SPE-DOM, with assigned cross peaks for compounds 2,4-di-tert-butylphenol (**M1**, $C_{14}H_{22}O$; red) and 3,3'-(5-(tert-butyl)-1,3-phenylene)

dipropionic acid (**M2**, $C_{16}H_{22}O_4$; blue); with ^{13}C-DEPT-135 (blue) and single pulse ^{13}C NMR spectra (black). The ^{1}H NMR resonance at $\delta_H \sim 7.31$ ppm (asterisk) shows no cross peaks and likely denotes a slowly exchanging acidic proton in a sterically hindered position. Green dotted lines are provided for better assessment of other main compounds present with structural features similar to **M1** and **M2**. ^{13}C NMR spectra at Fig. S16D and S16E are ^{13}C single pulse (black) and CH_2-selective (blue) ^{13}C DEPT NMR spectra.

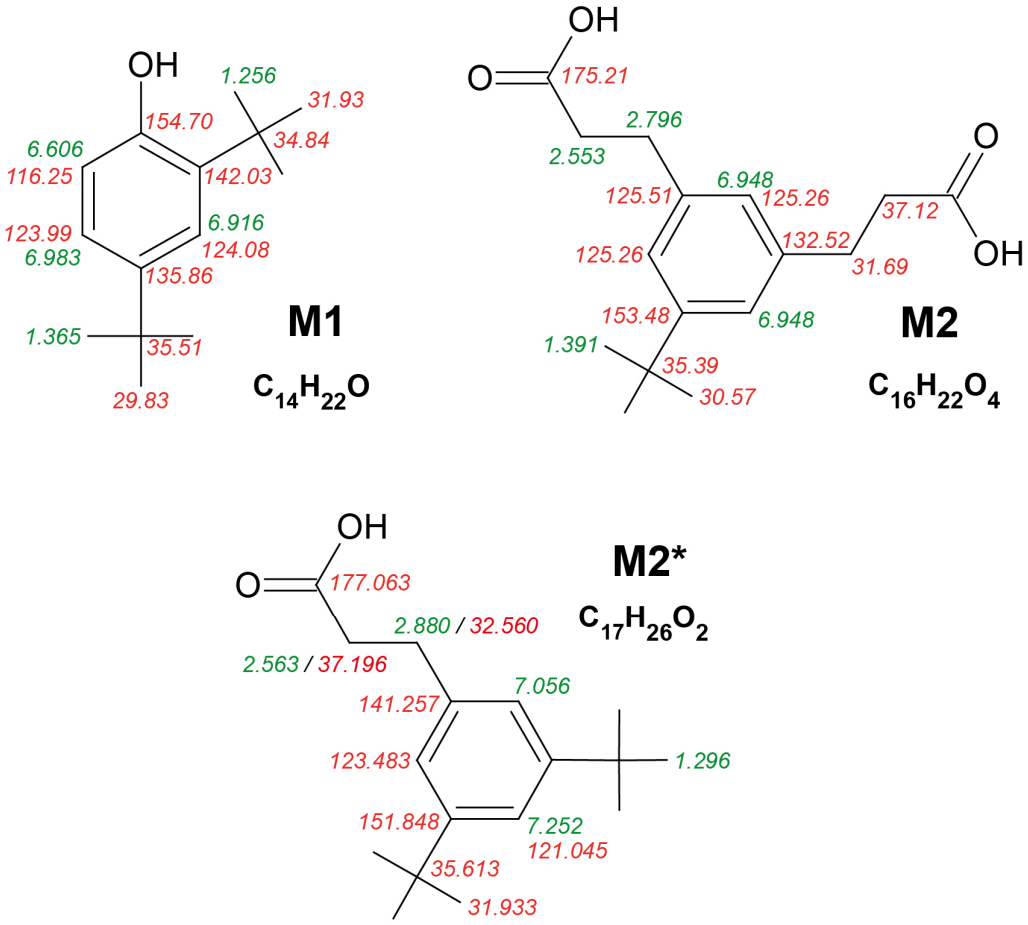

Fig. S17. ^{1}H (green) and ^{13}C NMR (red) chemical shifts of compounds **M1** (2,4-di-tert-butylphenol, $C_{14}H_{22}O$) and **M2** 3,3'-(3-tert-butylbenzene-1,5-diyl)dipropanoic acid, $C_{16}H_{22}O_4$), and commercially available compound M2* (3-(3,5-Di-tert-butylphenyl)propionic acid, $C_{17}H_{26}O_2$), which was absent in SCS SPE-DOM L05_1000m.

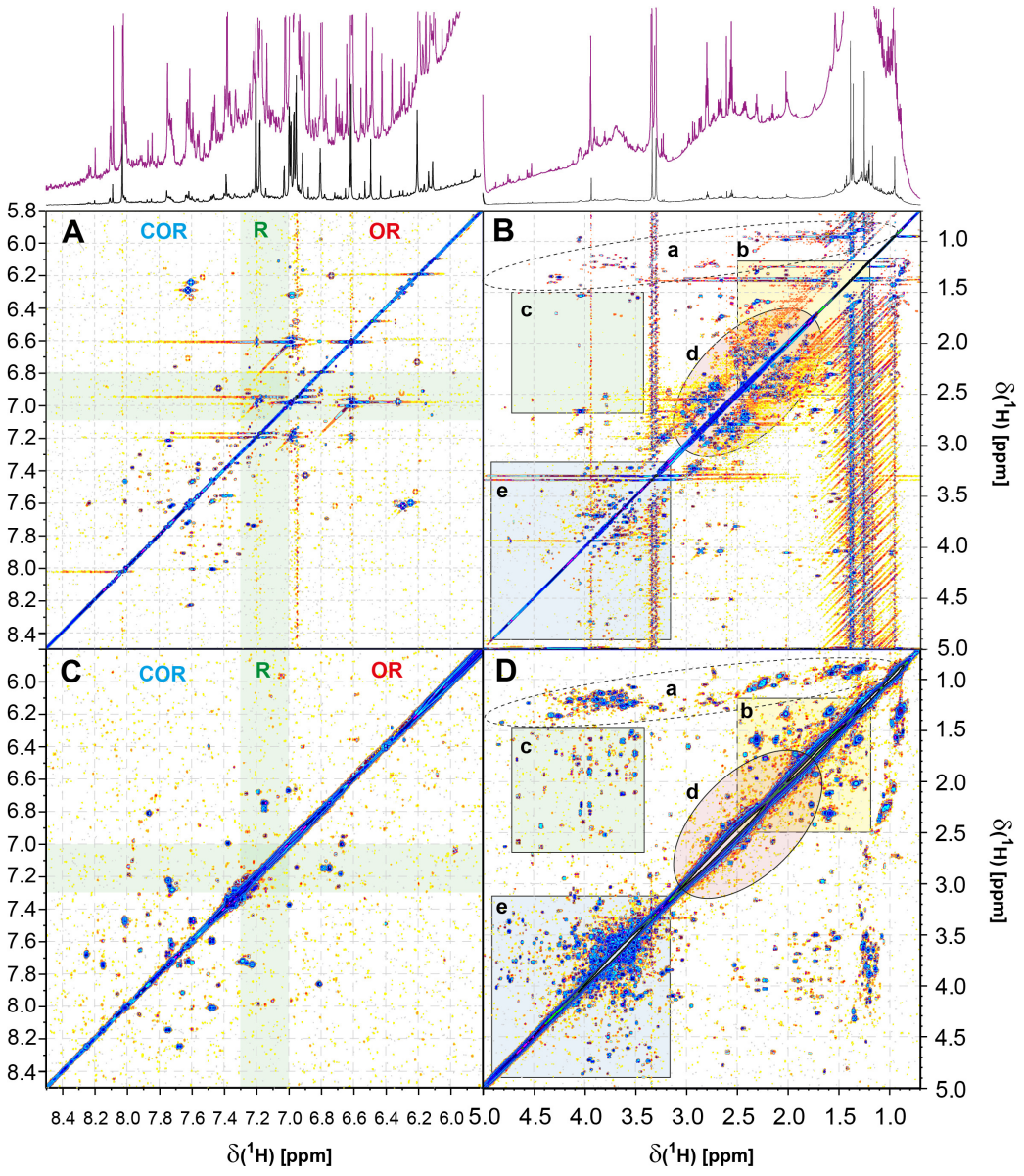

Fig. S18. ^{1}H, ^{1}H TOCSY NMR spectrum (15 msec mixing time) of SCS L05_1000 m SPE-DOM (**A**, **B)** and ^{1}H NMR spectrum above, and ^{1}H, ^{1}H COSY NMR spectrum of Atlantic Ocean fluorescence maximum zone SPE-DOM at 48m (panels **C**, **D**) (Hertkorn et al. 2013) with key substructures indicated. (**A**, **C)**: aromatic section with key substituents provided: electron withdrawing, polycyclic aromatic hydrocarbons and six-membered N-heterocycles (δ_H > 7.3 ppm), neutral (δ_H ~ 7-7.3 ppm), electron donating, olefins and five membered heterocycles (δ_H < 7 ppm), according to (Perdue et al. 2007). (**B**, **D)**: aliphatic section. section a: **H**$_3$CC**H** cross peaks related to chain terminating methyl groups, for δ_H > 3 ppm with oxygenated carbon (**H**$_3$CC**H**-O-); section b: intra aliphatic (X)HC-**H**C-**H**C-CH(Y) cross peaks excluding methyl; section c: functionalized aliphatics connected with oxygenated carbon units Z-C-**H**C-C**H**-O [Z = O (N, C$_{ar}$)]; section d: -**H**Cβ-Cα**H**-COX cross peaks with carbonyl derivatives, in case of L05_1000m SPE-DOM multiply substituted carboxylic acids; section e: -O-**H**CC**H**-O- cross peaks, carbohydrates and esters, ethers and alcohols.

References

Amaral, V., Graeber, D., Calliari, D. and Alonso, C. (2016) Strong linkages between DOM optical properties and main clades of aquatic bacteria. Limnology Oceanography 61(3), 906-918.

Catalá, T.S., Álvarez-Salgado, X.A., Otero, J., Iuculano, F., Companys, B., Horstkotte, B., Romera-Castillo, C., Nieto-Cid, M., Latasa, M. and Moran, X.A.G. (2016) Drivers of fluorescent dissolved organic matter in the global epipelagic ocean. Limnology and Oceanography 61(3), 1101-1119.

Catalá, T.S., Reche, I., Fuentes-Lema, A., Romera-Castillo, C., Nieto-Cid, M., Ortega-Retuerta, E., Calvo, E., Alvarez, M., Marrasé, C. and Stedmon, C.A. (2015) Turnover time of fluorescent dissolved organic matter in the dark global ocean. Nature Communications 6, 5986.

Cawley, K.M., Butler, K.D., Aiken, G.R., Larsen, L.G., Huntington, T.G. and McKnight, D.M. (2012) Identifying fluorescent pulp mill effluent in the Gulf of Maine and its watershed. Marine Pollution Bulletin 64(8), 1678-1687.

Coble, P.G. (1996) Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. Marine chemistry 51(4), 325-346.

Coble, P.G.J.C.r. (2007) Marine optical biogeochemistry: the chemistry of ocean color. Chemical reviews 107(2), 402-418.

Gonçalves-Araujo, R., Granskog, M.A., Bracher, A., Azetsu-Scott, K., Dodd, P.A. and Stedmon, C.A. (2016) Using fluorescent dissolved organic matter to trace and distinguish the origin of Arctic surface waters. Scientific Reports 6, 33978.

Guéguen, C., Cuss, C.W., Cassels, C.J. and Carmack, E.C. (2014) Absorption and fluorescence of dissolved organic matter in the waters of the Canadian Arctic Archipelago, Baffin Bay, and the Labrador Sea. Journal of Geophysical Research: Oceans 119(3), 2034-2047.

Helms, J.R., Stubbins, A., Perdue, E.M., Green, N.W., Chen, H. and Mopper, K.J.M.C. (2013) Photochemical bleaching of oceanic dissolved organic matter and its effect on absorption spectral slope and fluorescence. Marine Chemistry 155, 81-91.

Hertkorn, N., Harir, M., Koch, B., Michalke, B. and Schmitt-Kopplin, P. (2013) High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. Biogeosciences 10, 1583-1624.

Kothawala, D.N., Murphy, K.R., Stedmon, C.A., Weyhenmeyer, G.A. and Tranvik, L.J. (2013) Inner filter correction of dissolved organic matter fluorescence. Limnology and Oceanography: Methods 11(12), 616-630.

Li, P., Chen, L., Zhang, W. and Huang, Q. (2015) Spatiotemporal distribution, sources, and photobleaching imprint of dissolved organic matter in the Yangtze estuary and its adjacent sea using fluorescence and parallel factor analysis. PloS one 10(6), e0130852.

Martínez-Pérez, A.M., Nieto-Cid, M., Osterholz, H., Catalá, T.S., Reche, I., Dittmar, T. and Álvarez-Salgado, X.A.J.S.R. (2017) Linking optical and molecular signatures of dissolved organic matter in the Mediterranean Sea. Scientific Report 7(1), 3436.

Murphy, K.R., Stedmon, C.A., Graeber, D. and Bro, R. (2013) Fluorescence spectroscopy and multi-way techniques. PARAFAC. Analytical Methods 5(23), 6557-6566.

Murphy, K.R., Bro, R. and Stedmon, C.A. (2014a) Chemometric analysis of organic matter fluorescence. Aquatic organic matter fluorescence, 339-375.

Murphy, K.R., Stedmon, C.A., Wenig, P. and Bro, R. (2014b) OpenFluor–an online spectral library of auto-fluorescence by organic compounds in the environment. Analytical Methods 6(3), 658-661.

Nelson, N.B. and Gauglitz, J.M. (2016) Optical signatures of dissolved organic matter transformation in the global ocean. Frontiers in Marine Science 2, 118.
Nelson, N.B., Siegel, D.A., Carlson, C.A. and Swan, C.M. (2010) Tracing global biogeochemical cycles and meridional overturning circulation using chromophoric dissolved organic matter. Geophysical Research Letters 37(3).
Osburn, C.L., Handsel, L.T., Peierls, B.L. and Paerl, H.W. (2016) Predicting sources of dissolved organic nitrogen to an estuary from an agro-urban coastal watershed. Environmental Science & Technology 50(16), 8473-8484.
Osburn, C.L., Mikan, M.P., Etheridge, J.R., Burchell, M.R. and Birgand, F. (2015) Seasonal variation in the quality of dissolved and particulate organic matter exchanged between a salt marsh and its adjacent estuary. Journal of Geophysical Research: Biogeosciences 120(7), 1430-1449.
Osburn, C.L. and Stedmon, C.A. (2011) Linking the chemical and optical properties of dissolved organic matter in the Baltic–North Sea transition zone to differentiate three allochthonous inputs. Marine Chemistry 126(1-4), 281-294.
Painter, S.C., Lapworth, D.J., Woodward, E.M.S., Kroeger, S., Evans, C.D., Mayor, D.J. and Sanders, R.J. (2018) Terrestrial dissolved organic matter distribution in the North Sea. Science of the Total Environment 630, 630-647.
Parlanti, E., Wörz, K., Geoffroy, L. and Lamotte, M. (2000) Dissolved organic matter fluorescence spectroscopy as a tool to estimate biological activity in a coastal zone submitted to anthropogenic inputs. Organic Geochemistry 31(12), 1765-1781.
Perdue, E.M., Hertkorn, N. and Kettrup, A. (2007) Substitution Patterns in Aromatic Rings by Increment Analysis. Model Development and Application to Natural Organic Matter. Analytical Chemistry 79(3), 1010-1021.
Podgorski, D.C., Zito, P., McGuire, J.T., Martinovic-Weigelt, D., Cozzarelli, I.M., Bekins, B.A. and Spencer, R.G. (2018) Examining Natural Attenuation and Acute Toxicity of Petroleum-Derived Dissolved Organic Matter with Optical Spectroscopy. Environmental Science & Technology 52(11), 6157-6166.
Reemtsma, T., These, A., Linscheid, M., Leenheer, J. and Spitzy, A. (2008) Molecular and Structural Characterization of Dissolved Organic Matter from the Deep Ocean by FTICR-MS, Including Hydrophilic Nitrogenous Organic Molecules. Environmental Science & Technology 42(5), 1430-1437.
Stedmon, C.A. and Bro, R. (2008) Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. Limnology and Oceanography: Methods 6(11), 572-579.
Stubbins, A., Lapierre, J.-F., Berggren, M., Prairie, Y.T., Dittmar, T. and del Giorgio, P.A. (2014) What's in an EEM? Molecular signatures associated with dissolved organic fluorescence in boreal Canada. Environmental Science & Technology 48(18), 10598-10606.
Tanaka, K., Kuma, K., Hamasaki, K. and Yamashita, Y. (2014) Accumulation of humic-like fluorescent dissolved organic matter in the Japan Sea. Scientific Reports 4, 5292.
Wagner, S., Jaffé, R., Cawley, K., Dittmar, T. and Stubbins, A. (2015) Associations between the molecular and optical properties of dissolved organic matter in the Florida Everglades, a model coastal wetland system. Frontiers in Chemistry 3, 66.
Walker, S.A., Amon, R.M., Stedmon, C., Duan, S. and Louchouarn, P. (2009) The use of PARAFAC modeling to trace terrestrial dissolved organic matter and fingerprint water masses in coastal Canadian Arctic surface waters. Journal of Geophysical Research: Biogeosciences 114(G4).

Weishaar, J.L., Aiken, G.R., Bergamaschi, B.A., Fram, M.S., Fujii, R. and Mopper, K. (2003) Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. Environmental Science & Technology 37(20), 4702-4708.
Yamashita, Y., Boyer, J.N. and Jaffé, R. (2013) Evaluating the distribution of terrestrial dissolved organic matter in a complex coastal ecosystem using fluorescence spectroscopy. Continental Shelf Research 66, 136-144.
Yamashita, Y., Panton, A., Mahaffey, C. and Jaffé, R. (2011) Assessing the spatial and temporal variability of dissolved organic matter in Liverpool Bay using excitation–emission matrix fluorescence and parallel factor analysis. Ocean Dynamics 61(5), 569-579.
Zepp, R.G., Sheldon, W.M. and Moran, M.A. (2004) Dissolved organic fluorophores in southeastern US coastal waters: correction method for eliminating Rayleigh and Raman scattering peaks in excitation–emission matrices. Marine Chemistry 89(1), 15-36.