Supporting Information

Water-soluble organic composition of the Arctic sea surface microlayer and association with ice nucleation ability

Rosie J. Chance[1], Jacqueline F. Hamilton[1], Lucy J. Carpenter[1], Sina C. Hackenberg[1], Stephen J. Andrews[1], Theodore W. Wilson[2*]

1. Wolfson Atmospheric Chemistry Laboratories, Department of Chemistry, University of York, Heslington, York, YO10 5DD, UK.
2. School of Earth and Environment, University of Leeds, Woodhouse Lane, Leeds, LS2 9TJ, UK.
*Now at Owlstone Medical Ltd., 162 Cambridge Science Park, Milton Road Cambridge, CB4 0GH, UK.

Number of pages - 23

Supplementary text sections - 1

Number of tables - 3

Number of figures - 8

FT-ICR-MS Formula assignment

Molecular formulae for all *m/z* peaks with signal to noise ratio greater than four were generated using the SmartFormula functionality within DataAnalysis 4.1 software (Bruker Daltonics, Bremen, Germany). In addition to unlimited C, H and O, the heteroatoms N (up to 4), S (up to four) and P (up to one) were allowed, and a formula error limit of 1 ppm was applied. This typically resulted in ~15000 different *m/z* peaks per sample, many of which had multiple possible suggested formulae. Only *m/z* values that satisfied the following criteria were considered further: (i) absent from the procedural extraction blank (at a signal to noise ratio of at least four); (ii) present in both analytical replicates; (iii) signal-to-noise ratio greater than ten. This reduced the number of *m/z* values to 1072 to 2022 per sample.

For each *m/z* peak, a unique formula was selected by applying the following criteria. Elemental ratios were limited to those fitting the following rules: $O/C \leq 1$, $0.3 \leq H/C \leq 2.25$, $N/C \leq 0.5$, $S/C \leq 0.2$, $P/C \leq 0.1$, $(S+P)/C \leq 0.2$, $N < O$, $S < O$, $O > (2S + P)$. Formulae with the following elemental combinations were also disregarded as being unlikely (after Hawkes et al., 2016 [1]): $N_{3-4}S_{1-4}$, $N_{3-4}P$, $N_{1-4}S_{3-4}$, $N_{1-4}S_{1-4}P$ and $S_{1-4}P$. Double bond equivalents (DBE; calculated as $DBE = 0.5(2C + N + P - H + 2)$) were required to be a non-negative integer value. Where application of these rules did not result in a unique molecular formula, the formula with the lowest number of heteroatoms was selected for bulk compositional analysis (van Krevelen plots and calculation of average O/C and H/C ratios). The mSigma value (a measure of how well the measured isotopic abundance fits the theoretical isotopic abundance for a given molecular formula) was also used to evaluate possible formulae in some cases, however this parameter was not available for all ions.

A total of 5044 different formulae were found in the seawater samples (n=7), 6154 in the microlayer samples (n=7) and 2127 in the boat blanks (n=2). On average, only 33 ± 5 % of ions with S/N >10 had formulae assigned to them. Ions with assigned formulae accounted for an average of 62 ± 3 % of the total summed ion intensity for each seawater sample, but 36 ± 5 % of the total summed ion intensity for SML samples.

There was little overlap between the formulae found in samples of the same type (Fig. S7), with 55% of seawater formulae being found in only one seawater sample. This is surprising given the high level of similarity between the seawater mass spectra (SI Figure S6), and may reflect ambiguity in formula assignment. Elemental composition information obtained by assigning formula to the FT-ICR-MS results is discussed here for completeness, but should be treated with caution. The overall distribution of formulae between different compound classes (e.g. CHO, CHON etc) was generally similar across sample types (Table S2). The distribution of compound classes and average elemental ratios were within the ranges reported in the literature (Table S2). The percentage (by number) of CHON compounds was significantly higher in microlayer samples than seawater samples (students t-test, $p < 1\%$), but was also high in the boat blanks suggesting this might be an artifact related to the sampling or sample volume.

H/C ratios were slightly higher for microlayer samples than seawater (students t-test, $p < 1\%$), with overall averages of 1.25 ± 0.05 and 1.09 ± 0.05 respectively, while DBE values were lower (students t-test, $p < 1\%$), with overall averages of 11 and 15 for SML and seawater respectively. However, for both average H/C ratio and DBE, boat blanks had values between those for seawater and microlayer, again suggesting the sampling differences could have contributed to the apparent difference. Average O/C ratios were similar across the different sample types (overall average of 0.4 for each; Table S2). Visualization of the molecular characteristics/assigned formula using van Krevelen plots showed the three sample types to be broadly similar (Fig S8). The most pronounced difference was the increased abundance of heteroatomic molecules with $H/C > 1.5$ and $O/C < 0.5$ in the microlayer samples compared to the seawater and boat blank samples.

For the more individual formula assignment of IN tracer ions, some possible formulae containing phosphate (PO_4) groups were generated. However, these formulae were considered unlikely because the fragmentation spectra of these ions did not contain ions of *m/z* 97 ($H_2PO_4^-$) or 79 (PO_3^-), which are characteristic of phosphate functional groups.

Table S1. Dates and positions for seawater (sw), microlayer (sml) and boat blank (bb) sample collection during cruise JR288.

Station	Date	Latitude, °N	Longitude, °E	Samples	Extracted vol., L
6	22/7/13	73.11	-13.10	sw	10
				sml	0.5
8	24/7/13	75.05	-8.81	sw	9.6
				sml	0.65+0.3 sw
10	26/7/13	76.24	-5.45	sw	10
				sml	1.1
11	27/7/13	78.00	-7.17	sw	10
				sml	0.68
12	28/7/13	78.89	-7.03	sml	0.675
12.5	29/7/13	77.45	-5.22	sw	10
				sml	0.735
13	30/7/13	77.42	3.44	sw	10
				sml	0.39
16	3/8/13	81.13	24.12	sml	1.0
17	4/8/13	83.32	33.73	sw	10
				sml	0.91
18	5/8/13	82.67	26.10	sw	10
				sml	0.775
				bb	0.775
19	6/8/13	81.00	34.83	sw	10
				sml	~1 (not recorded)
21	9/8/13	76.75	9.37	sw	10
				bb	1.0

Table S2. Molecular composition information of SPE-DOM, as determined by FT-ICR-MS in negative ion mode in this study and comparable previous work. See main text for details of data filtering and formula assignment. Intensity weighted average values are indicated by 'wa' subscript.

	% formula type						Elemental ratios				Saturation	
	CHO	CHON	CHOS	CHOP	CHONS	CHONP	H/C	H/C_{wa}	O/C	O/C_{wa}	DBE	DBE_{wa}
Seawater												
sw6	38	18	9	21	8	6	1.03	1.12	0.36	0.41	16.0	12.2
sw8	38	18	8	20	9	6	1.09	1.15	0.36	0.43	15.5	11.9
sw12.5	38	18	10	21	8	6	1.02	1.11	0.37	0.43	16.4	12.4
sw13	48	19	4	14	8	7	1.17	1.2	0.43	0.49	12.5	10.3
sw17	41	20	10	17	8	4	1.13	1.17	0.39	0.43	12.6	11.0
sw18	39	17	8	25	8	4	1.09	1.16	0.38	0.45	14.9	11.3
sw19	35	22	9	19	10	6	1.09	1.14	0.37	0.4	15.2	12.1
average ± stdev	40 ± 4	19 ± 2	8 ± 2	20 ± 3	8 ± 1	5 ± 1	1.09 ± 0.05	1.15 ± 0.03	0.38 ± 0.02	0.43 ± 0.03	14.7 ± 1.6	11.6 ± 0.8
Microlayer												
sml6	37	22	13	13	11	4	1.19	1.05	0.42	0.43	11.3	14.5
sml8	30	22	11	21	10	5	1.2	1.37	0.38	0.44	11.9	9.5
sml12	36	24	8	20	8	5	1.28	1.33	0.39	0.39	10.9	10.1
sml12.5	42	25	8	16	6	4	1.2	1.22	0.44	0.46	10.4	9.6
sml17	33	23	8	15	15	4	1.28	1.28	0.37	0.36	11.1	11.0
sml18	34	24	10	17	10	6	1.24	1.23	0.39	0.41	10.8	10.3
sml19	30	26	11	19	10	5	1.33	1.34	0.37	0.36	10.3	11.3
average ± stdev	35 ± 4	24 ± 1	10 ± 2	17 ± 3	10 ± 3	5 ± 1	1.25 ± 0.05	1.26 ± 0.1	0.39 ± 0.03	0.41 ± 0.04	10.97 ± 0.6	10.9 ± 1.7
Boat blanks												
bb18	38	23	10	15	7	8	1.13	1.17	0.4	0.45	13.3	11.0
bb21	35	21	12	14	10	7	1.14	1.18	0.36	0.40	14.1	11.5
average	37	22	11	14	8	7	1.1	1.2	0.38	0.42	13.7	11.3

Selected literature values:												
L2013												
- sw	44.1	33.1	17.4	--	5.4	--	--	1.253	--	0.492	--	9.26
-sml	42.5	33.7	19.4	--	4.5	--	--	1.266	--	0.484	--	9.00
L2015												
- sw	43	30	4	< 2	2	19	--	1.2	--	0.5	--	--
- ice	39	39	4	< 3	2	13	--	1.4	--	0.4	--	--
H2013												
-sw (5 m)	24.6	30.8	21.4	--	23.3	--	--	1.6	--	0.54	--	--
K2009												
- sw	84	4	10.5	--	1.4	--	--	1.3	--	0.35	--	9.2

L2013 = Lechtenfeld, 2013[2]; L2015 = Longnecker, 2015[3]; H2013 = Hertkorn et al., 2013[4]; K2009 = Kujawinski et al., 2009[5].

Table S3. Sampling information and presence or absence of ions with the same molecular formulae as the IN tracer ions in ambient Arctic aerosol samples collected in the sampling area in spring and summer. Where ions were present, the formulae error in ppm is given, followed by the mSigma score.

	Sample collection				Tracer ion				
Sample	Date range	Latitude range, °N	Longitude range, °E	Air mass origin	$C_{10}H_{19}O_3^-$	$C_{11}H_{21}O_3^-$	$C_{20}H_{40}O_8S_2^{2-}$	$C_{12}H_{21}O_6^-$	$C_{14}H_{25}O_6^-$
March 2013									
RVL 2-5	18/3/2013 - 22/3/2013	77.5 - 80.3	-4.1 - 15.6	Barents sea	Absent	Absent	Absent	Absent	Absent
RVL 8-11	24/3/2013 - 28/3/2013	70.5 - 75.0	-16.9 - -7.0	north Greenland	Absent	Absent	Absent	Absent	Absent
RVL 01	16/3/2013 - 17/3/2013	n/a	n/a	Blank	Absent	Absent	Absent	Absent	Absent
RVL 15	31/3/2013 - 1/4/2013	n/a	n/a	Blank	Absent	Absent	Absent	Absent	Absent
July-August 2013									
JR288 5-7	17/7/2013 - 20/7/2013	65.9 - 70.2	-20.7 - -12.5	east Greenland	-2.61, 12.5	2.53, 8.8	Absent	-1.78, 28.2	-1.09, 519.5
JR288 20-23	2/8/2013 - 6/8/2013	79.8 - 83.3	10.7 - 34.8	north Greenland	3.27, 13.2	-3.37, 7.7	Absent	-2.3, 81.3	-1.81, 32
JR288 01	14/7/2013 - 15/7/2013	n/a	n/a	Blank	Absent	Absent	Absent	Absent	Absent
JR288 02	14/7/2013	n/a	n/a	Blank	Absent	Absent	Absent	Absent	Absent

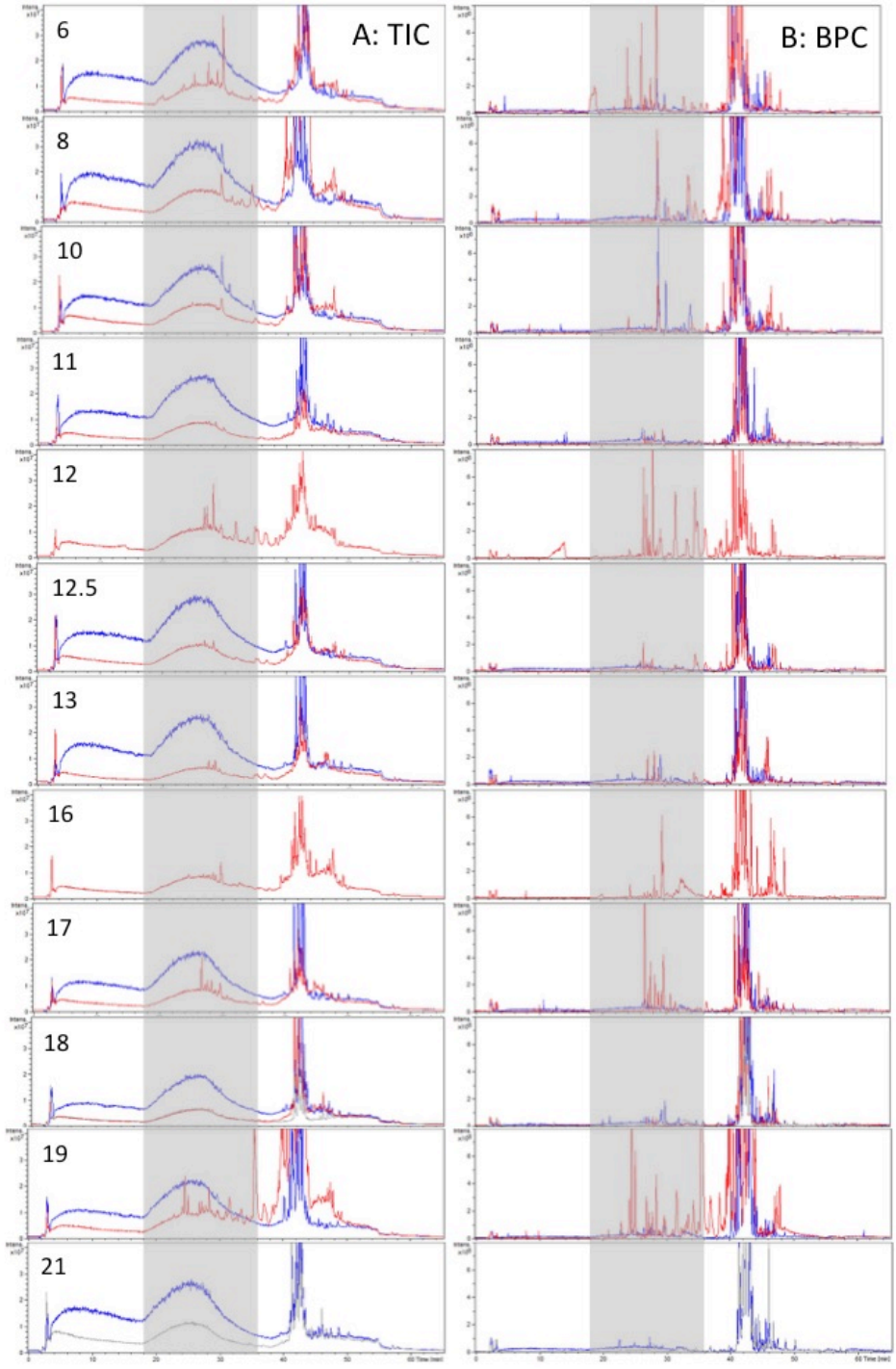


Figure S1. Negative mode total ion (A, left) and base peak (B, right) chromatograms for seawater (blue), microlayer (red) and microlayer sampler 'boat blank' (grey) sample extracts, obtained using LC-IT-MS. Station number for each pair of chromatograms shown at left. Light grey box shows the retention time range over which average mass spectra were calculated.

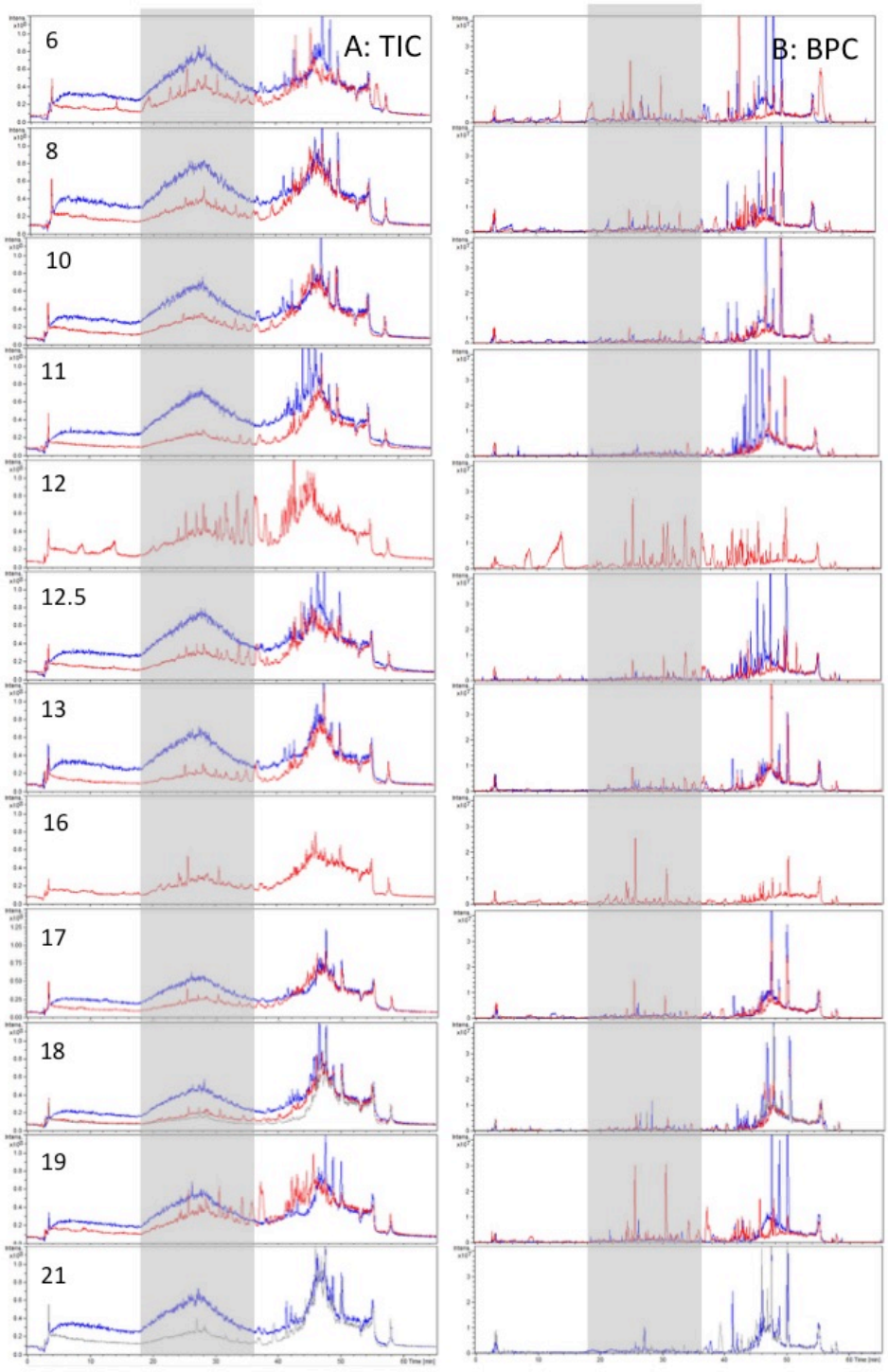
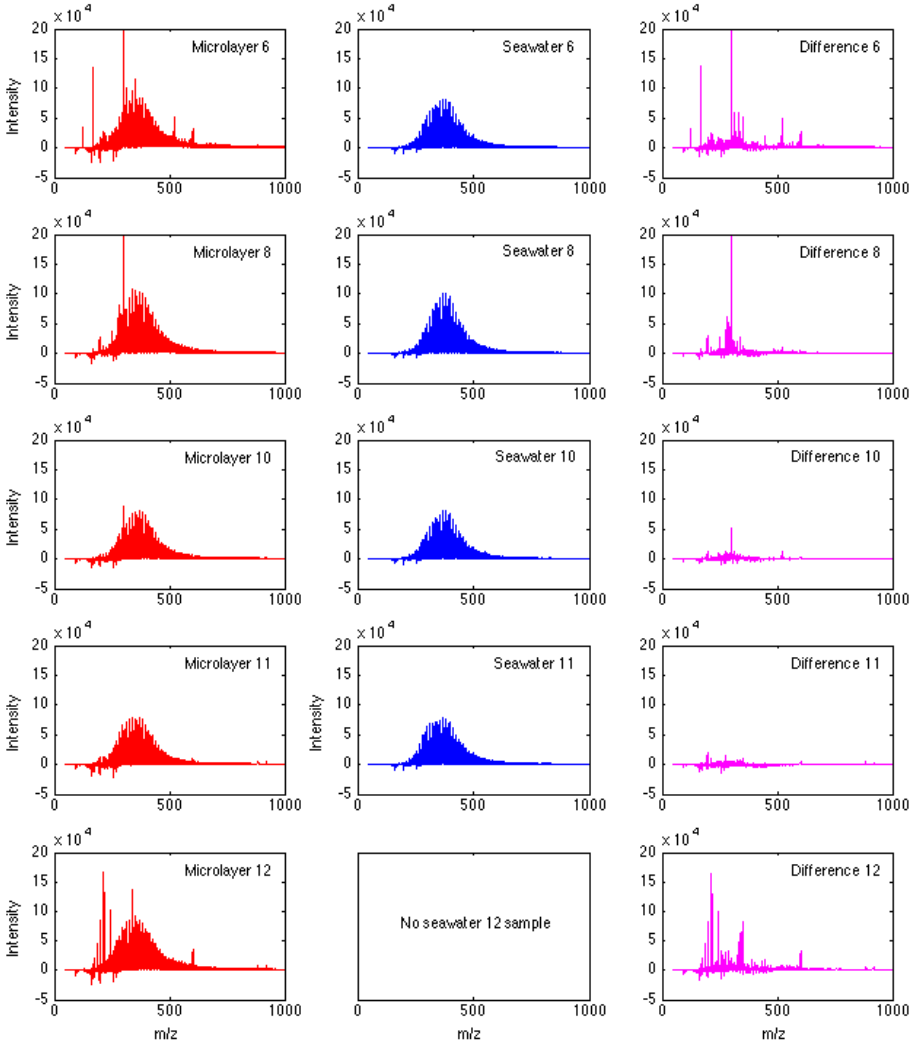
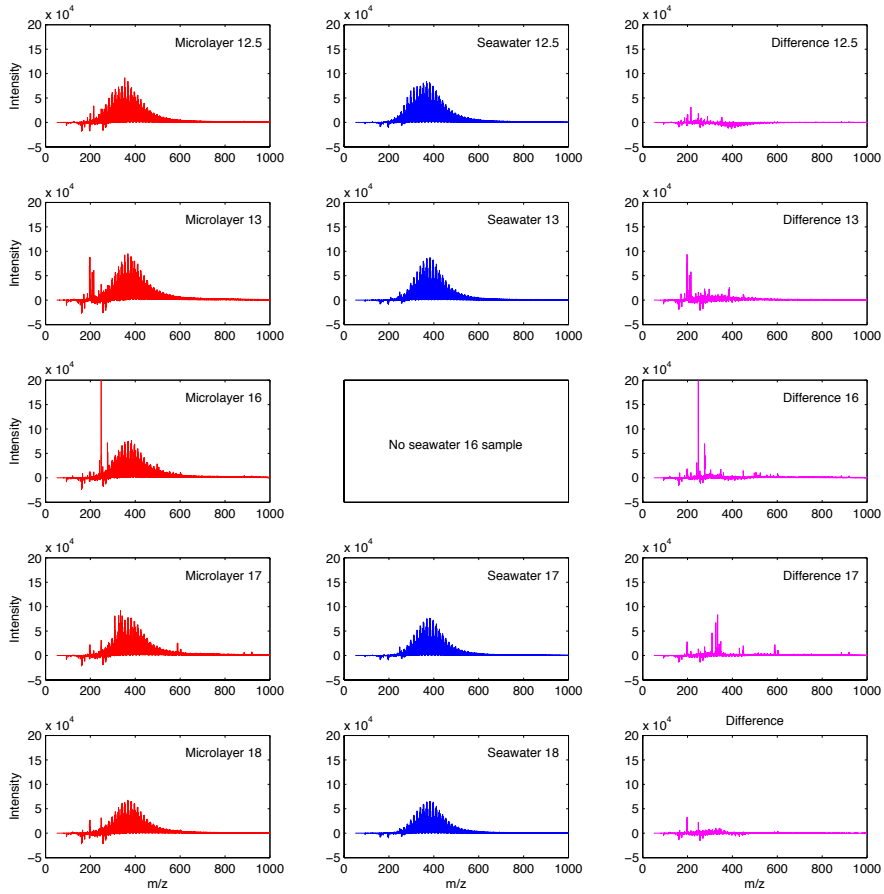


Figure S2. Positive mode total ion (A, left) and base peak (B, right) chromatograms for seawater (blue), microlayer (red) and microlayer sampler 'boat blank' (grey) sample extracts, obtained using LC-IT-MS. Station number for each pair of chromatograms shown at left. Light grey box shows the retention time range over which average mass spectra were calculated.


Microlayer 6
Seawater 6
Difference 6
Microlayer 8
Seawater 8
Difference 8
Microlayer 10
Seawater 10
Difference 10
Microlayer 11
Seawater 11
Difference 11
Microlayer 12
No seawater 12 sample
Difference 12
Intensity
m/z
x 10^4


Microlayer 12.5
Seawater 12.5
Difference 12.5
Microlayer 13
Seawater 13
Difference 13
Microlayer 16
No seawater 16 sample
Difference 16
Microlayer 17
Seawater 17
Difference 17
Microlayer 18
Seawater 18
Difference
Intensity
m/z
x 10^4

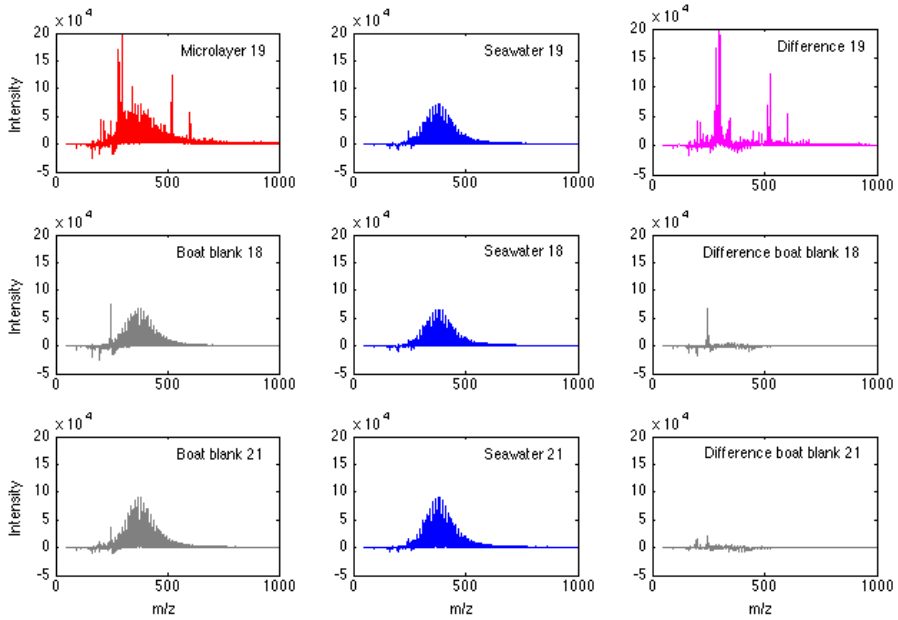
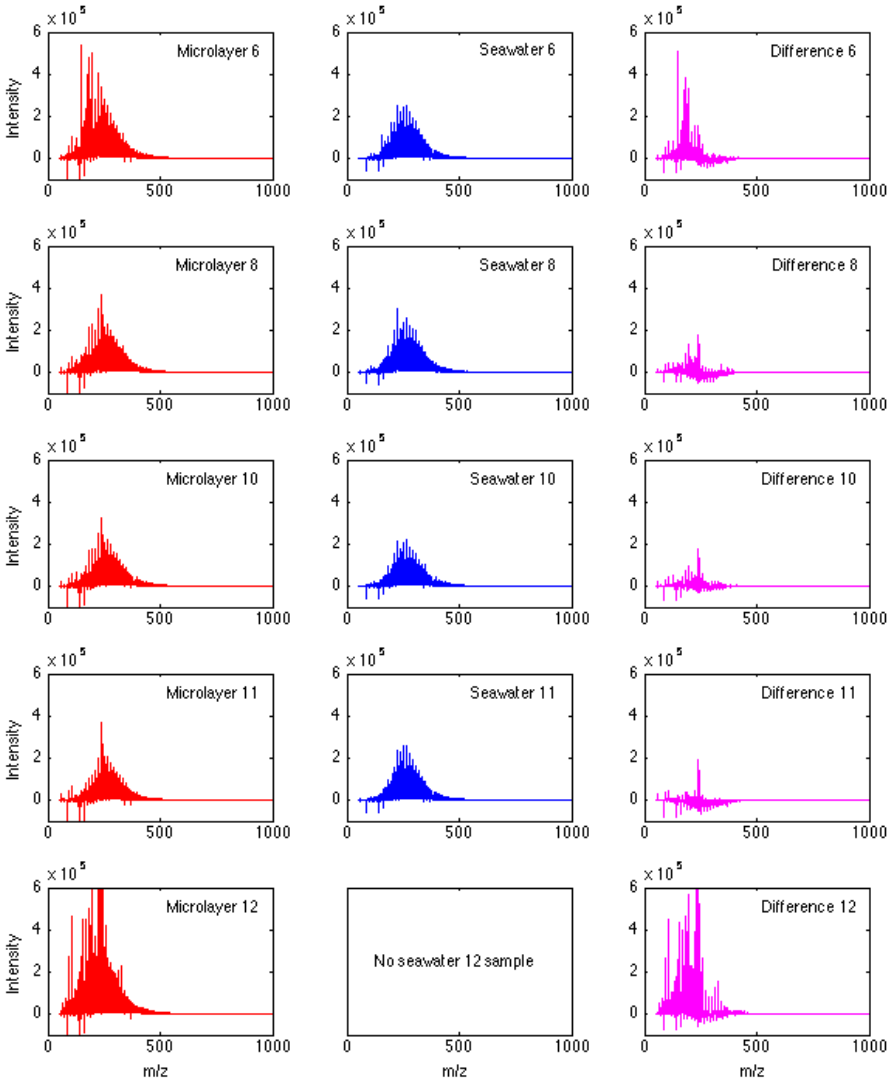
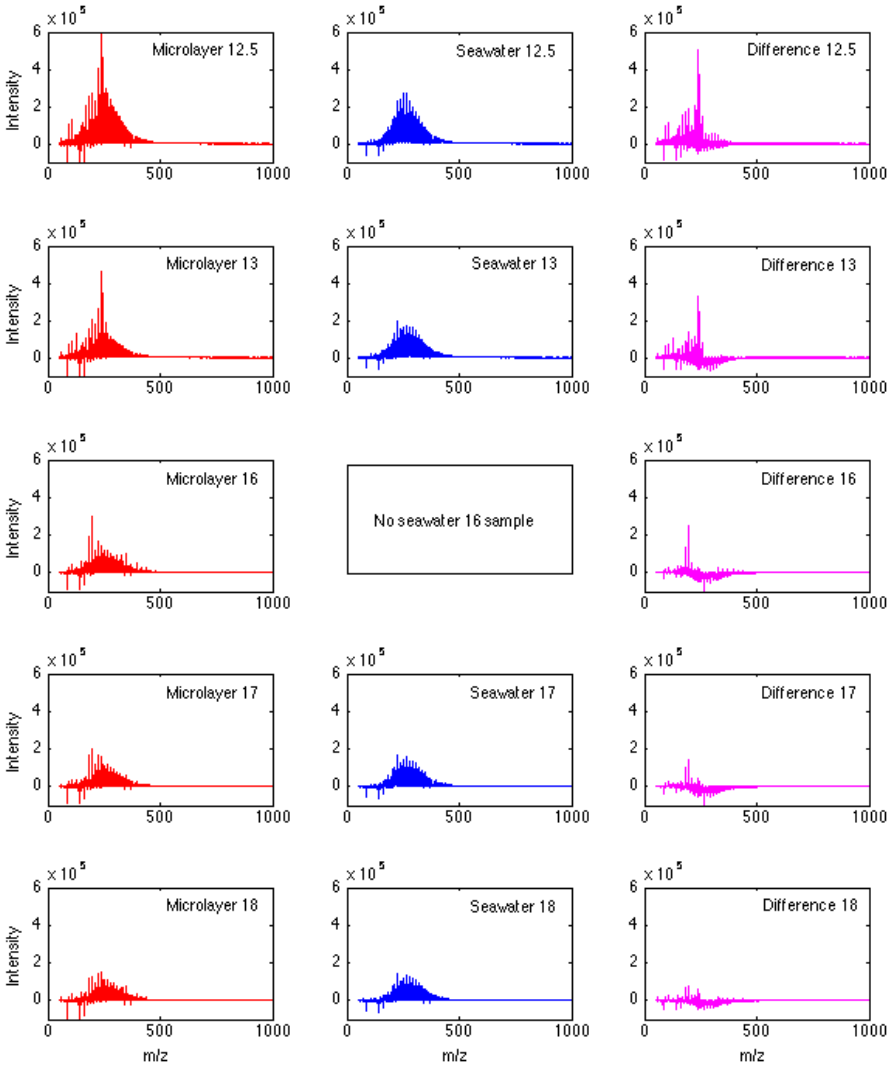

Microlayer 19
Seawater 19
Difference 19
Boat blank 18
Seawater 18
Difference boat blank 18
Boat blank 21
Seawater 21
Difference boat blank 21
Intensity
m/z
x 10^4

Figure S3A. Negative mode average mass spectra for retention time window 18 to 36 minutes for microlayer (red, left) and seawater (blue, center) samples, and normalized difference spectra (magenta, right) for these sample pairs. Spectra are normalized to total ion count, in order to account for the different sample loadings, and corrected for ions present in the extraction blank.


Microlayer 6
Seawater 6
Difference 6
Microlayer 8
Seawater 8
Difference 8
Microlayer 10
Seawater 10
Difference 10
Microlayer 11
Seawater 11
Difference 11
Microlayer 12
No seawater 12 sample
Difference 12
Intensity
m/z
x 10^5


Microlayer 12.5
Seawater 12.5
Difference 12.5
Microlayer 13
Seawater 13
Difference 13
Microlayer 16
No seawater 16 sample
Difference 16
Microlayer 17
Seawater 17
Difference 17
Microlayer 18
Seawater 18
Difference 18
Intensity
m/z
x 10^5

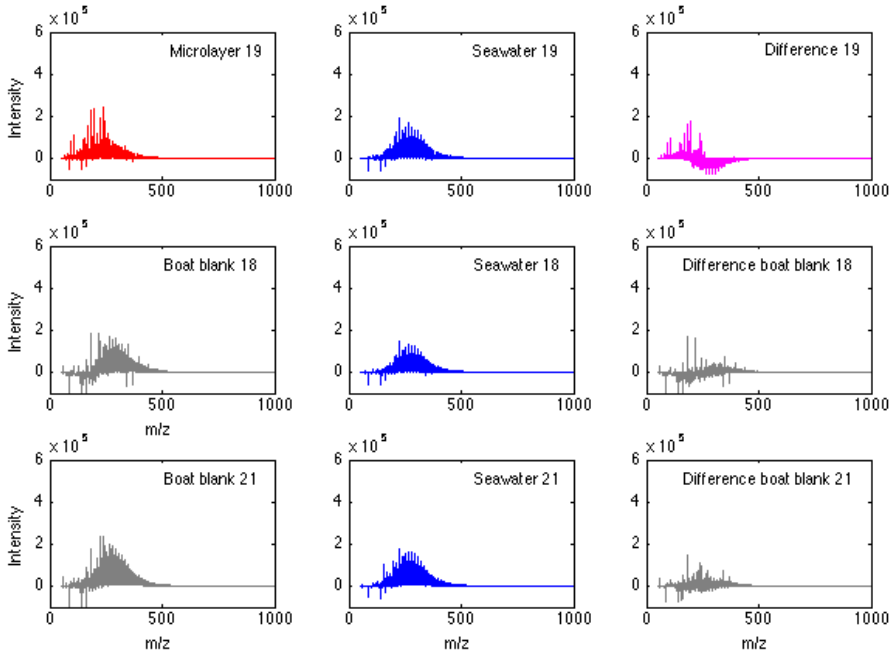


Figure S3B. Positive mode average mass spectra for retention time window 18 to 36 minutes for microlayer (red, left) and seawater (blue, center) samples, and normalized difference spectra (magenta, right) for these sample pairs. Spectra are normalized to total ion count, in order to account for the different sample loadings, and corrected for ions present in the extraction blank.

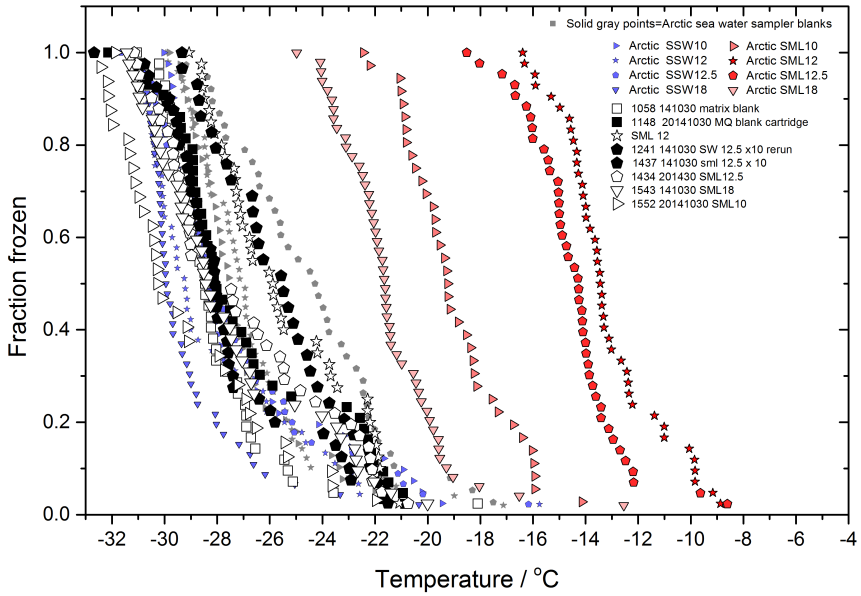


Figure S4. Freezing curves for microlayer extracts reconstituted in salt-water matrix (black and white symbols), and raw microlayer (red symbols) and seawater (blue symbols) samples assayed immediately following collection.

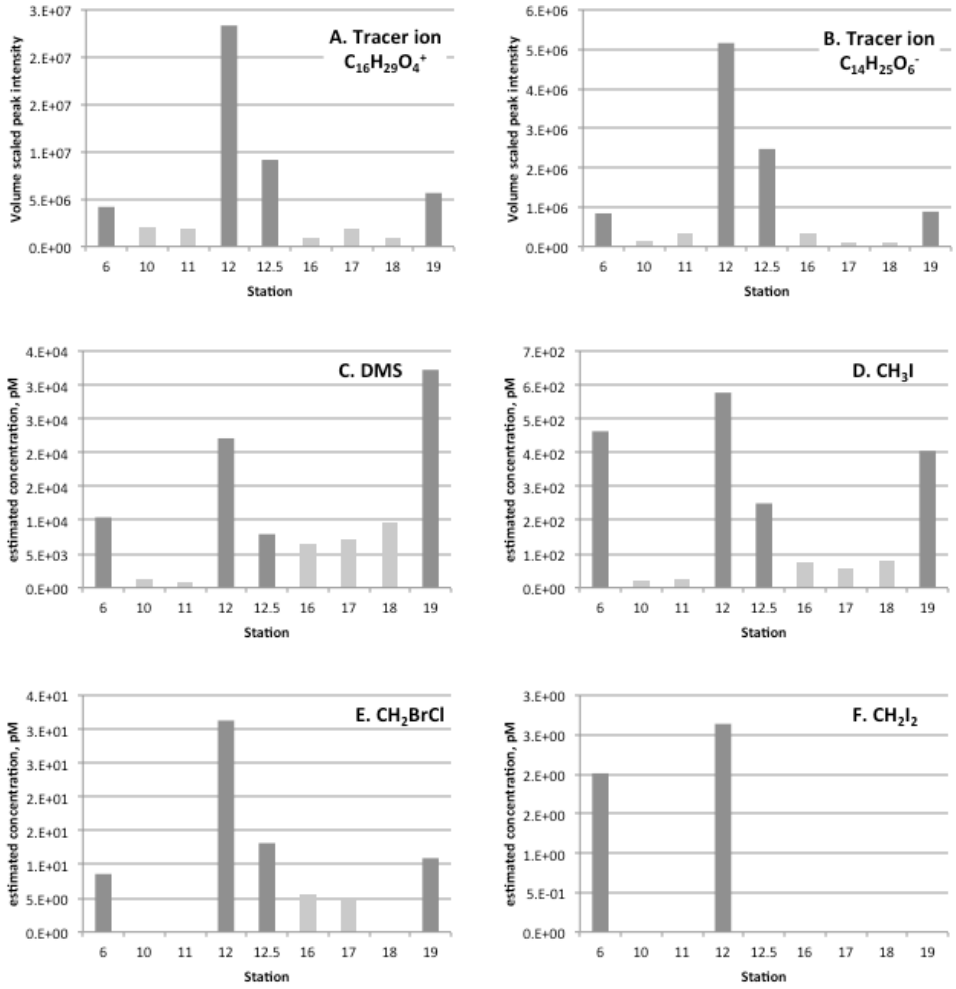
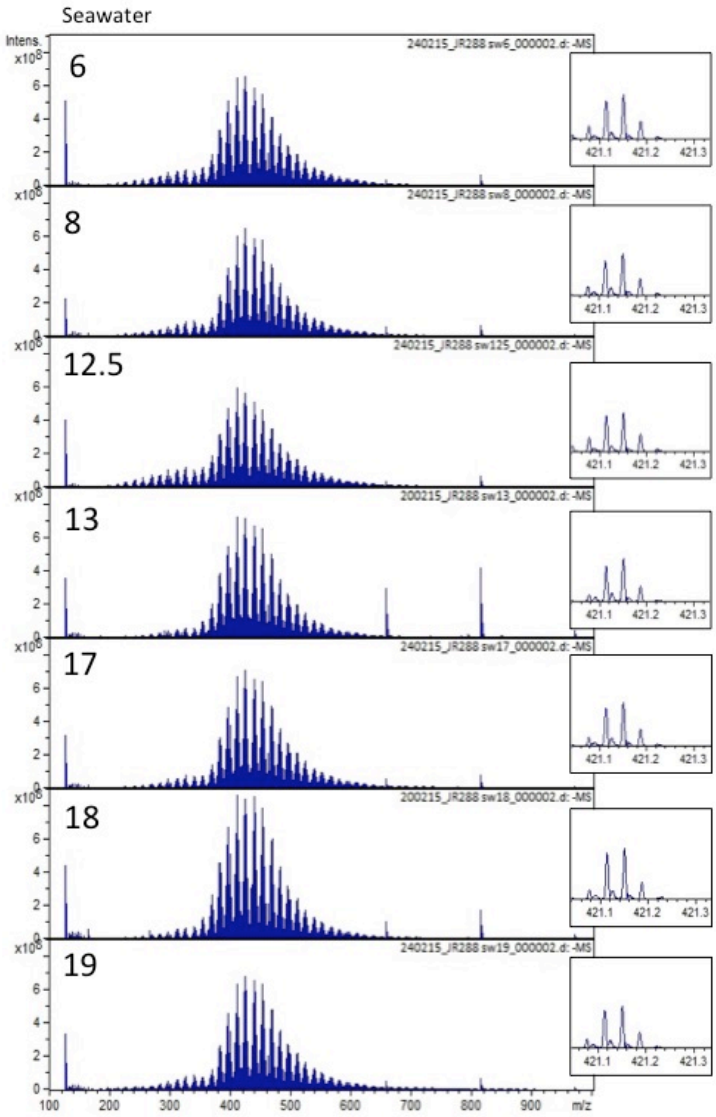
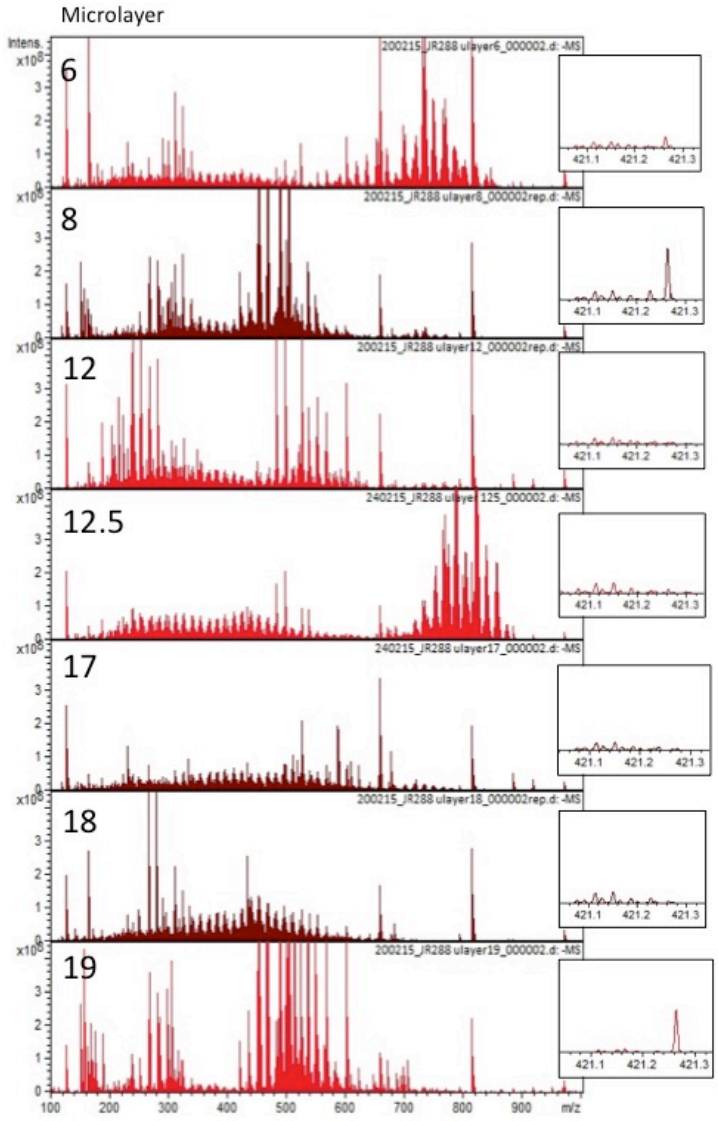


Figure S5. Bar charts showing relative abundance of the positive and negatively charged tracer ions most strongly correlated with IN activity (A. *m/z* 289, RT 30.8 mins; B. *m/z* 285, RT 38.5 mins), and the approximate concentrations of selected trace gases (C. Dimethyl sulfide; D. Iodomethane; E. Bromochloromethane; F. Diiodomethane) in microlayer samples.


Seawater
Intens.
x10^8
6
8
12.5
13
17
18
19
240215_iR288 sw6_000002.d: -MS
240215_iR288 sw8_000002.d: -MS
240215_iR288 sw125_000002.d: -MS
200215_iR288 sw13_000002.d: -MS
240215_iR288 sw17_000002.d: -MS
200215_iR288 sw18_000002.d: -MS
240215_iR288 sw19_000002.d: -MS
421.1
421.2
421.3
100
200
300
400
500
600
700
800
900
m/z


Microlayer
Intens.
x10^8
6
8
12
12.5
17
18
19
421.1
421.2
421.3
100
200
300
400
500
600
700
800
900
m/z

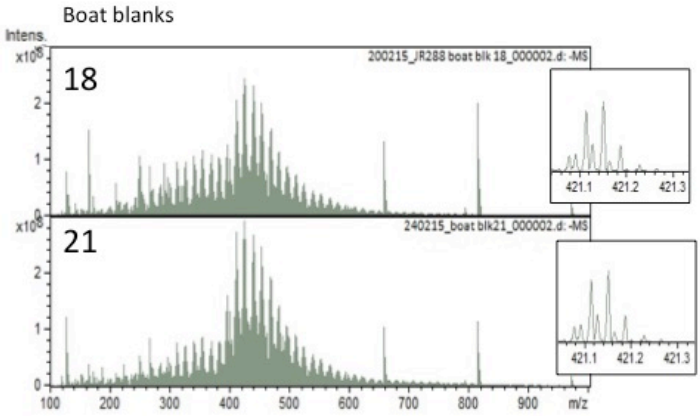


Figure S6. Negative mode FT-ICR-MS spectra for seawater (blue; top to bottom: 6, 8, 12.5, 13, 17, 18, 19), microlayer (red; top to bottom: 6, 8, 12, 12.5, 17, 18, 19) and microlayer sampler blank SPE extracts (grey; top to bottom: 18, 21). Inset panels show the fine scale resolution (across an arbitrary *m/z* range of 421.0-421.3) using the same vertical scale. The three most prominent seawater peaks were identified as $C_{20}H_{21}O_{10}$ (*m/z* = 421.12), $C_{21}H_{25}O_9$ (*m/z* = 421.15) and $C_{22}H_{29}O_8$ (*m/z* = 421.19) in all samples. SML samples with higher IN activity are shown in bright red, while those with lower IN activity are dark red

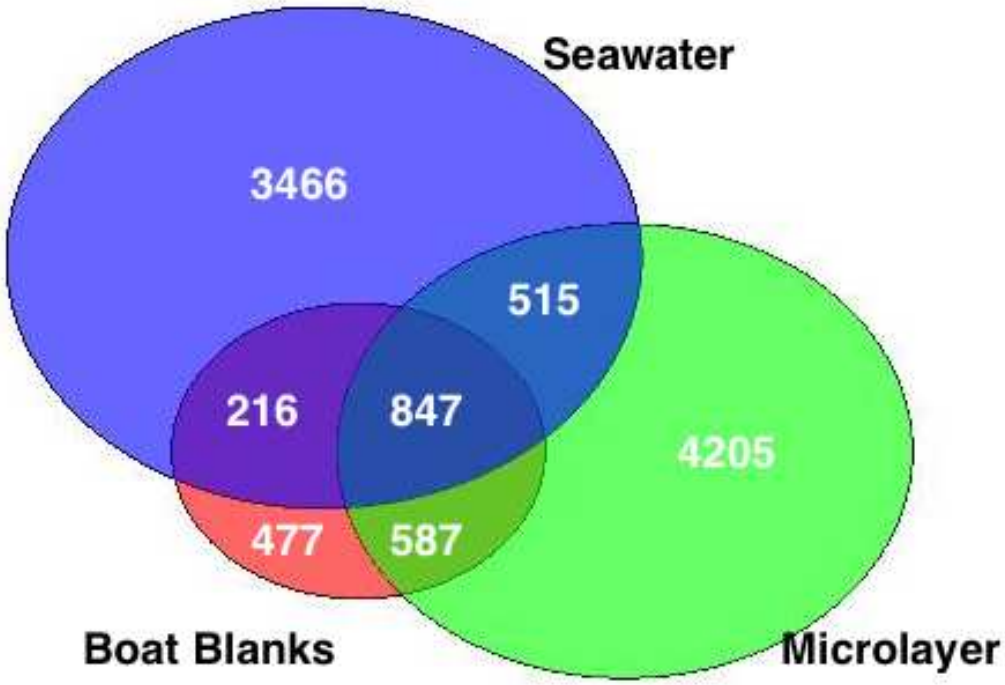


Figure S7. Venn diagram showing numbers of molecular formulae found for each sub-set of samples using FT-ICR-MS in negative ion mode.

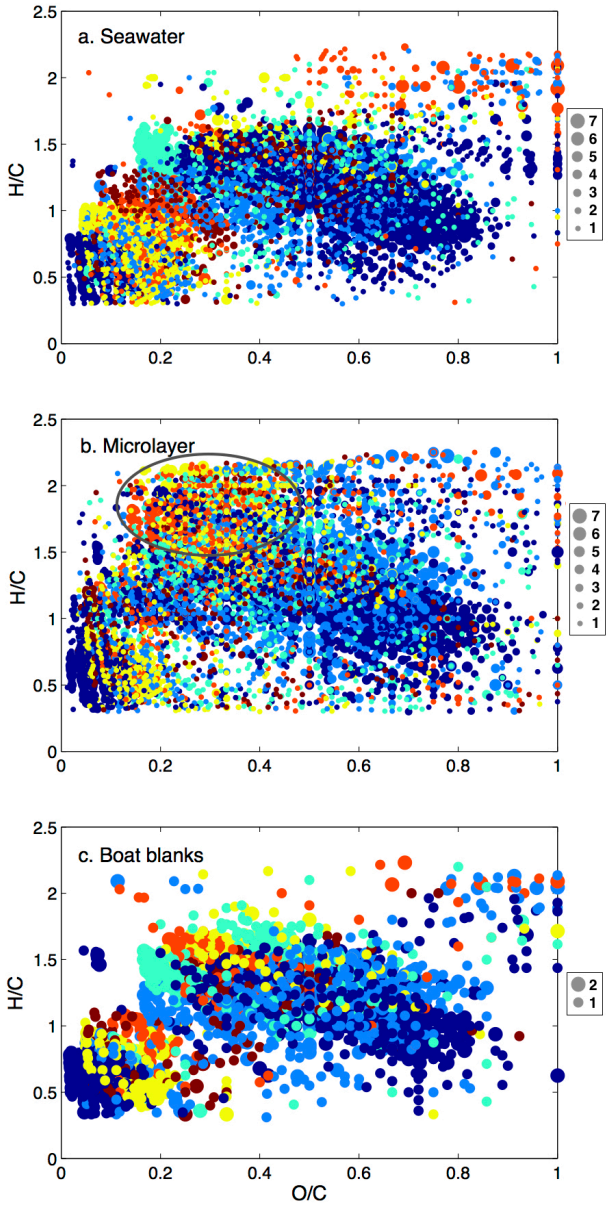


Figure S8. Van Krevelen plots for formulae identified in negative FT-ICR-MS spectra for (a) all seawater samples analyzed (n=7), (b) all microlayer samples analyzed (n=7) and (c) both sampling boat blanks (n=2). The size of the dots is scaled to the fraction of samples each formulae was found in. The color indicates the elemental composition as follows: dark blue = CHO, light blue = CHON, cyan = CHOS, yellow = CHOP, orange = CHONS and burgundy = CHONP. Grey oval highlights region where microlayer shows particular enrichment (see text for details).

References

1. Hawkes, J. A.; Hansen, C. T.; Goldhammer, T.; Bach, W.; Dittmar, T. Molecular alteration of marine dissolved organic matter under experimental hydrothermal conditions. *Geochim. Cosmochim. Acta* **2016,** *175*, 68-85.

2. Lechtenfeld, O. J.; Koch, B. P.; Gasparovic, B.; Frka, S.; Witt, M.; Kattner, G. The influence of salinity on the molecular and optical properties of surface microlayers in a karstic estuary. *Mar. Chem.* **2013,** *150*, 25-38.

3. Longnecker, K. Dissolved organic matter in newly formed sea ice and surface seawater. *Geochim. Cosmochim. Acta* **2015,** *171*, 39-49.

4. Hertkorn, N.; Harir, M.; Koch, B. P.; Michalke, B.; Schmitt-Kopplin, P. High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosciences* **2013,** *10* (3), 1583-1624.

5. Kujawinski, E. B.; Longnecker, K.; Blough, N. V.; Del Vecchio, R.; Finlay, L.; Kitner, J. B.; Giovannoni, S. J. Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochim. Cosmochim. Acta* **2009,** *73* (15), 4384-4399.