1 Illuminating the deep: Molecular signatures of photochemical alteration of dissolved organic
2 matter from North Atlantic Deep Water

3 Aron Stubbins^{a*}, Thorsten Dittmar^b

4
5 ^aSkidaway Institute of Oceanography, Marine Sciences Department, University of Georgia. 10
6 Ocean Science Circle, Savannah, Georgia, 31411, U.S.A.

7
8 ^bResearch Group for Marine Geochemistry (ICBM-MPI Bridging group), Institute for Chemistry
9 and Biology of the Marine Environment (ICBM), Carl von Ossietzky University of Oldenburg,
10 D-26111 Oldenburg, Germany. Telephone: +49-4441-798-3602; Fax: +49-441-798-3404. Email:
11 thorsten.dittmar@uni-oldenburg.de

12
13 * corresponding author:

14 aron.stubbins@skio.uga.edu

15 +1-912-598-2320 (office)

18 ***Keywords***

19 Dissolved organic carbon, photo-degradation, North Atlantic Deep Water, FTICR-MS, aromatic

20

21 *Abbreviations*

22 a , Napierian absorption coefficient; AI_{mod} , modified aromaticity index; BPCA,
23 benzenepolycarboxylic acid; CDOM, chromophoric dissolved organic matter; CHO, molecular
24 formula containing only Carbon, Hydrogen and Oxygen; CHON, molecular formulas containing
25 C, H, O and Nitrogen; CHOS, molecular formulas containing C, H, O and Sulfur; CHOP,
26 molecular formulas containing C, H, O and Phosphorous; CHONS, molecular formulas
27 containing C, H, O, N and S; Da, Daltons; DBC, dissolved black carbon; DOC, dissolved
28 organic carbon; DOM, dissolved organic matter; ESI, electrospray ionization; FTICR-MS,
29 Fourier transform ion cyclotron mass spectrometry; NADW, North Atlantic Deep Water; UV,
30 ultraviolet.

31

32

33

34

35

36

37

38

39

40

41 *Abstract*

42 Duplicate, filtered samples of North Atlantic Deep Water (NADW) were irradiated for 28
43 days in a solar simulator. Duplicate dark controls were placed alongside the irradiated samples.
44 After 28 days, samples were extensively photo-degraded based upon colored dissolved organic
45 matter (CDOM) photo-bleaching (>95%). Samples were solid phase extracted using PPL resin to
46 isolate, concentrate and desalt the dissolved organic matter (DOM) in the samples. Ultrahigh
47 resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry
48 (FTICR-MS) enabled 3024 molecular formulas to be identified in the dark control. Photo-
49 degradation decreased molecular diversity, with 2402 formulas assigned post-irradiation.
50 Molecular formulas were classified based upon their photo-lability as 1) photo-resistant; 2)
51 photo-labile; and, 3) photo-produced. Photo-resistant DOM made up 73% of all formulas and
52 was dominated by highly unsaturated molecular signatures consistent with carboxylic-rich
53 alicyclic molecules, suggesting these apparently bio-refractory compounds may also survive
54 multiple passages through sunlit surface waters and thus accumulate for timeframes exceeding
55 ocean ventilation. Photo-labile DOM was enriched in low molecular weight formulas, aromatic
56 molecular formulas, and sulfur and phosphorous containing formulas. Compounds containing
57 both sulfur and nitrogen were particularly photo-labile and may represent an underappreciated
58 component of the photo-reactive CDOM pool. Photo-produced DOM was enriched in higher
59 molecular weight formulas, as well as aliphatic and peptide formulas. Molecular formulas are
60 indexed by their photo-lability classification in the supplementary information.

61 *Introduction*

62 Oceanic dissolved organic carbon (DOC) constitutes a major pool of global carbon,
63 approximately equivalent in size to the atmospheric load of carbon dioxide (Dittmar and
64 Stubbins, 2014). Thus, the longterm stability of oceanic DOC is of direct consequence for the
65 radiative balance of the earth. The majority of the oceanic store of DOC is contained in the deep
66 ocean (Hansell, 2013). Plots of DOC versus salinity in the deep ocean indicate that net removal
67 of DOC is extremely slow, leading to the definition of the deep ocean DOC pool as refractory
68 (Hansell, 2013). This view of extreme recalcitrance in the deep ocean is supported by bulk
69 radiocarbon mixing models, which suggest the refractory DOC in the deep ocean has a turnover
70 time of around 6,000 years (Bauer, 2002). More recent radiocarbon dating of operationally
71 defined fractions of marine DOC has revealed that bulk (i.e. average) apparent DOC ages do not
72 represent the average value for a normally distributed pool of molecules (Follett et al., 2014;
73 Walker et al., 2011). Follett et al. (2014) radiocarbon dated DOC fractions produced during a
74 time series of ultraviolet (UV) irradiance. Decay and mixing models fitted to the generated data
75 indicate the presence of a highly radiocarbon depleted pool of DOC (-1000 to -600 $\delta^{14}\text{C}$) that
76 occurs at elevated concentrations in the deep ocean and has an apparent turnover time of 30,000
77 years, assuming no external sources of radiocarbon depleted DOC exist (Follett et al., 2014).

78 In the darkness of the abyss, it appears a fraction of ancient carbon persists. However, in
79 surface waters this ancient carbon is actively cycled. For instance, the most radiocarbon depleted
80 fraction identified by Follett et al. (2014) occurs at significantly lower concentrations in the
81 surface (9 μM) than in the deep (24 μM), suggesting a surface removal process is at play. Other
82 research has shown that radiocarbon depleted carbon is utilized by surface microbes (Cherrier et
83 al., 1999). Both the quantitative loss (Follett et al., 2014) and the increased bioavailability

(Cherrier et al., 1999) of deep ocean refractory DOM once it enters surface waters may be facilitated by photochemical processes.

In sunbathed surface waters, DOC is bombarded by photons. When light is absorbed by the colored fraction of dissolved organic matter (CDOM), a multitude of photoreactions ensue (Mopper et al., 2015). Some DOC is mineralized directly to inorganic forms such as carbon dioxide or carbon monoxide (Powers and Miller, 2015; Stubbins et al., 2006). The fraction of DOC that survives photo-degradation is altered (Mopper et al., 2015). Depending upon the nature of the original DOC and irradiation conditions, photo-alteration can increase the bioavailability of DOC (Mopper et al., 2015). The photo-generation of bioavailable DOC from refractory DOC has been invoked to explain the utilization of ancient carbon by surface microbes mentioned in the previous paragraph (Cherrier et al., 1999). Thus, photochemistry plays a direct (photo-mineralization to dissolved inorganic carbon) and biologically mediated (photo-production of biolabile substrates) role in the removal of DOC from sunlit waters. In addition to these removal processes, photochemistry leaves behind a fraction of DOM with modified optical, bulk and molecular composition (Gonsior et al., 2009; Kujawinski et al., 2004; Medeiros et al., 2015; Osburn et al., 2001; Spencer et al., 2009; Stubbins et al., 2010). In the current paper, we focus upon the photo-alteration of deep ocean DOM molecular signatures for a sample collected from North Atlantic Deep Water (NADW).

The photo-lability (i.e. the ease with which a substrate can be degraded) of DOM has been related to optical and molecular properties. For instance, CDOM is preferentially degraded during photo-irradiation experiments (Helms et al., 2013; Helms et al., 2008). The light absorbing properties of CDOM stem from aromatic moieties within the DOM pool (Stubbins et al., 2008; Weishaar et al., 2003), suggesting aromatic signatures should be preferentially photo-

degraded relative to non-aromatic ones. Chemical characterization of photo-irradiated DOM bears this out. Multiple lines of evidence, for samples of multiple origin (e.g. terrigenous and marine) and utilizing multiple chemical analyses (e.g. nuclear magnetic resonance spectroscopy, specific compound analysis, and high resolution mass spectrometry), all reveal the preferential photo-degradation of aromatic moieties (Helms et al., 2014; Spencer et al., 2009; Stubbins et al., 2012).

Fourier transform ion cyclotron mass spectrometry (FTICR-MS) brings a major fraction of DOM into our analytical window, providing high accuracy molecular-level information regarding elemental and inferred structural composition that can be related to source (Dittmar and Stubbins, 2014; Kujawinski et al., 2002; Mopper et al., 2007) and has been used to address the photochemical transformations of DOM (Gonsior et al., 2009; Kujawinski et al., 2004; Medeiros et al., 2015; Stubbins et al., 2010). Here we apply FTICR-MS to elucidate the molecular signatures of DOM photo-alteration for a sample of NADW. This sample was previously analyzed for DOC, CDOM and dissolved black carbon photo-degradation (Stubbins et al., 2012). In the current study, FTICR-MS data was used to determine the molecular signatures of the following three photochemically defined pools of DOM and to contrast the nature of these pools with the same fractions observed in Congo River water DOM (Stubbins et al., 2010):

- 1) Photo-resistant DOM: molecular signatures that survive irradiation;
- 2) Photo-labile DOM: molecular signatures are were lost during irradiation, and;
- 3) Photo-produced DOM: molecular signatures that are formed during irradiation.

Materials and methods

Sample Collection

Bermuda Atlantic Time Series (BATS) cruise 252 on board the RV Atlantic Explorer (6th to 10th November 2009) was joined to collect North Atlantic Deep Water (NADW), one of the major global deep ocean water masses. Water masses and their dissolved organic constituent loads undergo modification during transport. For NADW at BATS, DOC concentrations are lower than in NADW near points of deep water formation around Greenland, indicating DOC entrained within NADW at deep water formation sites slowly biodegrades or is otherwise removed as the water mass transits to our sampling site near Bermuda (Hansell and Carlson, 1998). The NADW sample was collected by CTD on the 9th of November 2009 from 3000 m at the BATS site (31°40' N; 64°10' W). Three Niskin bottles were fired at 3000 m. As soon as the CTD was on deck and secure, large surface area 0.2 µm capsule filters were attached (Polycap TC, Whatman) directly to the Niskin bottles' nipples and sample was gravity filtered into a 20 L fluorinated high density polyethylene carboy (Nalgene). The filters and carboy were precleaned by soaking for at least 24 hours in acidified ultrapure (MilliQ) water (pH 2 with hydrochloric acid; p.a.), soaking in basic ultrapure water for a further 24 hours (0.1 M sodium hydroxide; p.a.), and rinsing with copious volumes of ultrapure water (filters flushed with >20 L; carboy quintuple rinsed). The sample was then stored frozen in the dark at -20°C onboard RV Atlantic Explorer until she docked in Norfolk, Virginia, USA (19/12/2009). At this point the carboy was collected and taken to Old Dominion University, approximately 10 minutes' drive from the dock, and returned to a -20°C freezer. From there, the sample was later transported the nine hours to Skidaway Institute of Oceanography (SkIO), Georgia, USA in a large cooler with other frozen samples and placed back in a -20°C freezer upon arrival. Thus, the sample remained frozen from shortly after collection until defrosting for the irradiation study.

Irradiations

All analyses, from the irradiation through the extraction and FTICR-MS analysis were conducted in duplicate. On the 20th of September 2010 the thawed sample was transferred to two 2 L precombusted spherical quartz irradiation flasks. A further 4 L aliquot was divided between two 2 L combusted borosilicate flasks, which were then wrapped in aluminum foil. All samples were then placed under a solar simulator fitted with 12 UVA-340 bulbs (integrated irradiance $\sim 25 \text{ W m}^{-2}$; Q-Panel), which provide a spectral shape and flux closely approximating natural sunlight from 295 to 365 nm (Stubbins et al., 2008), the main wavelength range for environmental photochemical reactions involving CDOM. The temperature in the solar simulator was kept at approximately 20°C using a side mounted fan. Sample water temperatures were likely higher as measured temperatures for test samples have been between 25°C and 30°C. Duplicate 2 L flasks were taken from the light table after 28 days. One day of irradiation using this solar simulator design is approximate to 1.27 times daily solar irradiance during the winter at 36.89°N or 0.67 times the daily (12 hour) irradiance at the equator (Helms et al., 2008; Spencer et al., 2009).

Fourier transform ion cyclotron resonance mass spectrometry

Duplicate samples were solid phase extracted using PPL Bond Elut (Agilent) resins before analysis by electrospray ionization FTICR-MS in negative ion mode. As reported in Stubbins et al. (2012), 1507 to 1975 mL of acidified NADW seawater was solid phase extracted following the method of Dittmar et al. (2008). The DOC extraction efficiency ranged from 67 to 74% (Stubbins et al., 2012). An ultrapure water sample was also loaded onto PPL, eluted and used as a procedural blank for extraction. The duplicate methanol extracts were diluted 1:1 with ultrapure water and analyzed in negative mode electrospray ionization using a 15 Tesla FTICR-

MS (Bruker Solarix) at the University of Oldenburg, Germany. 500 broadband scans were accumulated for the mass spectra. After internal calibration, mass accuracies were within an error of < 0.2 ppm. Molecular formulas were assigned to peaks with signal to noise ratios greater than five based on published rules (Koch et al., 2007; Singer et al., 2012; Stubbins et al., 2010). Peaks detected in the procedural blank (PPL extracted ultrapure water) were removed. Peak detection limits were standardized between samples by adjusting the dynamic range of each sample to that of the sample with the lowest dynamic range (dynamic range = average of the largest 20% of peaks assigned a formula divided by the signal to noise threshold intensity; standardized detection limit = average of largest 20% of peaks assigned a formula within a sample divided by the lowest dynamic range within the sample set) (Spencer et al., 2014; Stubbins et al., 2014). Peaks below the standardized detection limit were removed to prevent false negatives for the occurrence of a formula within samples with low dynamic range.

Assigned formulas were categorized by compound class based upon elemental stoichiometries (Stubbins et al., 2010). Modified aromaticity index (AImod) (Koch and Dittmar, 2006) values were calculated:

$$\text{AImod} = (1 + \text{C} - 0.5\text{O} - \text{S} - 0.5\text{H}) / (\text{C} - 0.5\text{O} - \text{S} - \text{N} - \text{P}) \quad (2)$$

AImod values 0.5 to 0.67 and >0.67 were assigned as aromatic and condensed aromatic structures, respectively (Koch and Dittmar, 2006). Further compound classes were defined as follows: Highly Unsaturated = AImod <0.5, H/C <1.5; Aliphatics = H/C 1.5 to 2.0, O/C <0.9, N = 0; and; Peptide molecular formulas = H/C 1.5 to 2.0, O/C <0.9 and N >0. It should be noted that compounds identified as “peptides” have the molecular formulas of peptides, but their actual structure may differ.

The main goal of the current study was to determine the molecular signatures of photo-resistant, photo-labile, and photo-produced DOM defined as:

1) Photo-resistant DOM - formulas assigned to peaks that survived irradiation and are present in both the duplicate initial and one of the irradiated samples;

2) Photo-labile DOM - formulas assigned to peaks that were lost during irradiation (i.e. present in both the duplicate pre-irradiation samples, and absent from both the post-irradiation samples), and;

3) Photo-produced DOM - formulas assigned to peaks that were formed during irradiation (i.e. absent from both the pre-irradiation samples, and present in both the post-irradiation samples).

In defining these groupings, it should be remembered that electrospray ionization is selective, and that this selectivity results in one peak's intensity changing relative to increases or decreases in other peak intensities. Therefore, a compound that appeared to be produced during the irradiation may have been present in the sample all along, but emerged from the background as other peaks were removed during photo-degradation. Alternatively, a compound identified as labile, may also still be present at the end of the experiment, but may not show up if photo-products with high ionization efficiencies are photo-produced. To account for this, aliquots of the pre-irradiation and post-irradiation PPL extracts were mixed together 1:1 based upon DOC concentrations and analyzed by FTICR-MS.

Results

The high resolution and precision of the FTICR-MS enabled the assignment of exact molecular formulas to thousands of peaks (Figure 1; Table 1). Both the pre- and post-irradiation samples provided high quality spectra that followed the peak distribution anticipated for PPL

extracts of marine DOM (Figure 1). Observation of ^{13}C isotopic peaks at a distance 1.00335 m/z from the parent ions indicated that peaks were singularly charged (Kujawinski et al., 2002; Stenson et al., 2002).

Unique molecular formulas were assigned to 3024 FTICR-MS for the initial, pre-irradiation sample of NADW (Table 1). Aromatic (5.4%) and condensed aromatic (0.1%) contributed a modest proportion of these formulas (Table 1). Highly unsaturated molecular formulas were most abundant (83.9%), followed by aliphatics (9.4%) and peptides (1.3%). With respect to elemental compositions, 45.1% of NADW molecular formulas were CHO-only, 43.3% contained N, 11.4% contained S, and 2.6% contained P (Table 1). The average mass of formulas was 457 Daltons, and average Almod, H/C and O/C were 1.26, 0.47 and 0.25 respectively (Table 1).

The FTICR mass spectrum for post-irradiation NADW DOM retained a vast amount of molecular information (Figure 1). However, the number of peaks with formula assigned fell from 3024 to 2402, a reduction of 21% (Table 1). This net loss of molecular complexity during photo-alteration was driven by a loss of photo-labile molecular formulas (23%), which was not compensated for by the photo-production of new formulas (3%; Table 1 and 2; Figure 2A).

As noted in the methods, as the intensities of produced or degraded peaks within the DOM matrix change, so the intensities of other peaks can also be influenced. This could give rise to the apparent production or loss of formulas during an experiment. For instance, a peak, let's call it Peak A, present in the initial sample at an intensity just below the detection limit, might exceed the detection limit post-irradiation due simply to the loss of other peaks and a concomitant increase in the relative abundance of Peak A. The reverse also holds. If Peak B is present at just above the detection limit prior to irradiation, and a suite of compounds with high

capacity for drawing charge are photo-produced, Peak B could drop below the detection limit. In the current experiment, both photo-produced and photo-labile formulas were identified – i.e. peaks were lost and others emerged during the irradiation. To ascertain if these molecular formulas were indeed degraded and produced, a 1:1 mix of the pre- and post-irradiation sample was analyzed. Of the 3130 unique formulas assigned to peaks that occurred above the detection limit in the pre- and post-irradiation samples, 95% were also present at or above the same detection limit in the 1:1 mix sample. Of the 3024 formulas observed in the initial, pre-irradiated sample, 95% of formulas were also observed in the 1:1 mix; whereas 99% of the 2402 formulas observed post-irradiation were observed in the 1:1 mix. Grouping molecules by photo-lability, 99% of all photo-resistant formulas were observed in the 1:1 mix. This is hardly surprising, as these formulas were also observable in the two samples used to create the 1:1 mix. Ninety three percent of the 106 photo-products and 82% of the 728 photo-labile formulas were observable in the 1:1 mix. This suggests that there is a modest bias between the pre- and post-irradiation samples that may have obscured 6% of apparent photo-products in the initial sample, and 18% of photo-labile moieties in the post-irradiation sample. The 6% of potentially misassigned photo-products constitutes just 7 formulas, and is thus a very modest number. The number of formulas potentially misassigned as photo-labile represents 130 formulas, which is more substantial. However, the population of photo-labile formulas also present in the 1:1 mix is still very similar to that defined by simple presence absence between the samples in Table 1. For instance, 100% of condensed aromatic, 77% of aromatic, 17% of highly unsaturated, 20% of aliphatics, and 41% of peptide compounds are labile when only formulas observed in the 1:1 mix are considered; whereas, these values are 100%, 81%, 19%, 27% and 40% when all formulas are considered (Table 1). Given the modest adjustment to the data and no clear method to determine which

mode of analysis is most valid, for the remainder of the document the formulas presented were those present or absent in the pre- and post-irradiated samples (i.e. no consideration is given to their presence or absence in the 1:1 mix).

Although all condensed aromatic compounds were photo-labile, their minor contribution to the initial sample translated to a minor contribution to the total photo-labile DOM pool (2 formulas or 0.3% of all photo-labile DOM; Table 1 and 2). Aromatics were also highly photo-labile (81% removal) and contributed 18.5% to the total photo-labile formula count (Table 1 and 2). Highly unsaturated moieties were the least photo-labile DOM component (19% photo-labile). However, as these compounds constituted 84.9% of the initial formulas in NADW DOM, they also constituted the greatest percentage of formulas to the labile DOM pool (66.3%). Twenty seven percent of aliphatics and 40% of peptides were photo-labile (Table 1). Highly unsaturated and aliphatic compounds were the most photo-resistant moieties (Table 1). The photo-product pool was enriched in aliphatics and peptides relative to initial DOM, but depleted in all other DOM types (Table 2).

Based upon elemental composition, CHO-only compounds were the most photo-resistant (82%; Table 1) and constituted modestly to the photo-labile pool (26.1%; Table 2), but provided the greatest contribution to the photo-product pool (58.5%; Table 2). Consequently, the post-irradiation sample was enriched in CHO-only compounds relative to the initial sample (Table 2). The post-irradiation sample was depleted in N, P and S containing formulas compared to the pre-irradiation sample (Table 2). With respect to N, this is despite the majority of compounds containing N being photo-resistant (70%, Table 1). On balance, the loss of 359 N containing formulas was not compensated for by production (43 formulas; Table 1). This effect is even more apparent for S containing (211 photo-labile; 2 photo-produced) and P containing formulas

(38 photo-labile; 0 photo-produced). The average properties of molecular formulas showed only a minor variation between the pre- and post-irradiation samples, with AImod decreasing (0.25 to 0.24), and H/C and O/C both increasing (Table 1). The shift in AImod was driven by the selective removal of high AImod formulas (average AImod for photo-labile DOM = 0.27) and the production of low AImod formulas (average AImod for photo-products = 0.13; Table 1). Similarly, the shifts in H/C and O/C were due to photo-labile DOM having lower and the photo-products having higher values than the initial, pre-irradiated sample (Table 1).

Discussion

FTICR-MS analysis revealed NADW DOM isolates to contain a diverse suite of molecular formulas (Table 1; Figure 3). The distribution of these formulas in van Krevelen space (i.e. photo-labile and resistant formulas in Figure 2B) and between molecular classifications (Table 1) are consistent with previous studies of PPL extracts of oceanic DOM (Flerus et al., 2012; Hertkorn et al., 2006; Hertkorn et al., 2013). Most notably, the NADW spectrum was dominated by highly unsaturated formulas (Table 1). Compounds with these molecular properties are consistent with a number of dissolved organic moieties, including highly functionalized aromatics, such as lignin derived compounds, and carboxylic rich alicyclic molecules (CRAM) (Hertkorn et al., 2006). Lignin is an aromatic biomarker derived from vascular land plants (Meyers-Schulte and Hedges, 1986); whereas, CRAM is believed to be of microbial origin, to have a predominantly autochthonous source in the ocean, and be a dominant component of the DOM pool within the oceans (Hertkorn et al., 2006). Given the marine origin and low CDOM (Stubbins et al., 2012) of the initial NADW sample, it is likely that the majority of highly unsaturated formulas represent non-aromatic, CRAM-like moieties. However, lignin in solid phase extracts and the $\delta^{13}\text{C}$ of high molecular weight DOM, both signatures of DOM

source, indicate that terrigenous, aromatic DOM is exported to NADW from the Arctic Ocean (Hernes and Benner, 2006). Hernes and Benner (2006) reported that such terrestrial inputs constitute only 1.6% of deep water DOC at BATS. Therefore, while it cannot be ruled out that a fraction of the highly unsaturated compounds in NADW contain aromatic carbon of autochthonous, marine or heterochthonous, terrigenous origin, it is likely that the majority of these formulas represent CRAM-like moieties.

As reported previously, DOC concentration decreased $76\pm 8\%$ during the irradiation of NADW, falling from an initial $50.8\pm 0.4\text{ }\mu\text{M}$ to $12.3\pm 4.2\text{ }\mu\text{M}$ over 28 days (Stubbins et al., 2012). CDOM light absorption at 300 nm decreased $95\pm 5\%$ from $0.35\pm 0.02\text{ m}^{-1}$ to $0.02\pm 0.02\text{ m}^{-1}$ during the irradiation (Stubbins et al., 2012). That close to 100% of CDOM absorbance was photo-labile is consistent with results from other long-term irradiation studies (Helms et al., 2014; Helms et al., 2013; Spencer et al., 2015; Spencer et al., 2009). Dissolved black carbon (DBC), quantified as condensed aromatics isolated from seawater via PPL solid phase extraction and quantified as benzenepolycarboxylic acid (BPCA) oxidation products, fell from $1044\pm 164\text{ nM-C}$ to $55\pm 15\text{ nM-C}$, a 20-fold decrease in concentration during the course of the irradiation (Stubbins et al., 2012). These data are reported here for context. Comparison of losses of total DOC ($76\pm 8\%$) with the more modest loss of photo-labile molecular formulas (23%; Table 1) might appear contradictory. However, this result is not surprising as the complete loss of FTICR-MS formulas was likely also accompanied by reductions in the peak intensities of other formulas and, fundamentally, FTICR-MS is not quantitative.

The preferential loss of CDOM and DBC reported previously (Stubbins et al., 2012), indicate preferential photo-degradation of aromatic compounds. This is confirmed by the preferential loss of the molecular signatures for condensed aromatics (100% photo-labile) and

aromatics (81% photo-labile) as revealed by FTICR-MS analysis (Table 1). The condensed aromatics reported here are analogous to DBC as quantified by the BPCA method in Stubbins et al. (2012), confirming that DBC is highly photo-labile. Thus, photo-chemistry is likely a primary sink of DBC and other bio-refractory organics in the ocean (Stubbins et al., 2012). DBC molecules, whether quantified as BPCAs or characterized by FTICR-MS, should also find use as exemplary tracers of the relative photo-degradation state of DOM, both in the field and during photo-degradation experiments.

Although it was apparent aromatics were preferentially removed, only subtle changes occurred in the average molecular properties of NADW DOM following photo-degradation. For instance, average molecular weight shifted from 457 to 461 Da and the modified aromaticity index from 0.25 to 0.24 (Table 1). These shifts are modest compared to those observed for river water DOM. For instance, Congo River water DOM photo-degraded under the same conditions and to approximately the same degree based upon CDOM removal (~95% of CDOM loss) shifted average molecular weight from 424 to 409 Da (Spencer et al., 2009; Stubbins et al., 2010). Comparing differences in the photo-lability of river and marine DOM molecular formulas can provide insight as to the original chemistry of the samples. To further assess molecular controls upon photo-lability and allow comparison to Congo River DOM, three pools of NADW DOM were defined based upon their photoreactivity: photo-resistant DOM, photo-labile DOM, and photo-produced DOM (see methods for full definitions).

The populations of molecular formulas belonging to each group were separated by mass, with photo-labile formulas having lower and photo-products greater average molecular weights than the initial DOM and the photo-resistant DOM (Table 1; Figure 3). The high mass of photo-products is consistent with results for Congo River DOM. However, photo-labile Congo River

DOM was of higher mass than photo-resistant DOM (Stubbins et al., 2010). Photo-labile NADW DOM was enriched in N, S, and P containing formulas relative to initial DOM; whereas the photo-resistant and photo-produced DOM fractions were depleted in S and P, but of similar N content to the initial DOM (Table 1). The accentuated photo-lability of CHOS compounds identified by FTICR-MS has been noted for DOM in acid mine drainage, with oxidation to sulfate offered as an explanation for the loss of organic S (Herzprung et al., 2010). Closer examination revealed that of the 61 molecular formulas that contained both N and S (i.e. CHONS) in NADW, 59 were photo-labile. Model organic molecules, such as glutathione and cysteine, which have S and N in close proximity in their structures, are efficient photo-producers of carbonyl sulfide (COS) upon irradiation; whereas other S containing organics without N or with N at a greater distance from S in the molecule yielded little or no COS upon irradiation (Stubbins and Dittmar, 2015). Thus, photo-production of COS as well as sulfate may provide a photo-chemical sink for organic S. The drivers of CHONS photo-lability remain unclear, yet simple non-aromatic CHONS compounds such as cysteine absorb UV light (Johnson, 1971) and are photo-reactive in pure solution (Stubbins and Dittmar, 2015), indicating that they are primary photo-reactants and may represent a significant, overlooked component of marine CDOM. As FTICR-MS only provides elemental formulas with no additional information as to the isomeric arrangements of the elements, it is impossible to say whether the photo-labile CHONS molecular formula contained N and S in close proximity in their structures.

As with previous irradiation studies (Gonsior et al., 2009; Kujawinski et al., 2004; Medeiros et al., 2015; Stubbins et al., 2010), the aromatic content of the molecular formulas drove the main trends in the FTICR-MS data. To graphically represent this, the proportions of formulas that were photo-resistant, photo-labile or photo-produced were plotted in a bar graph

versus the modified aromaticity index (AImod; Figure 2B). This same plot was also used to represent the molecular signatures of Congo River DOM photo-lability (Stubbins et al., 2010). As for Congo River water DOM, all molecular formulas with an AImod greater than 0.6 and present in the initial samples were photo-labile. With decreasing AImod, the proportion of photo-labile molecular formulas decreased and the percentage of photo-resistant molecular formulas increased, the latter peaking at 0.3 to 0.4 (Figure 2B). The presence of a high number of photo-resistant molecular formulas with AImod values greater than 0.3 is at odds with the results for Congo River DOM, for which photo-resistant DOM peaked at AImod 0.1 to 0.2 (Stubbins et al., 2010). For NADW, the majority of photo-resistant formulas fell within the classification highly unsaturated (Table 1) and occupied the region of van Krevelen space (H/C 0.75 to 1.75; O/C 0.15 to 0.8; Figure 3) where many biogeochemically distinct isomers are possible for any given molecular formula. Of pertinence to photo-lability, the formulas could either be aromatic, possibly vascular land plant derived lignin containing compounds, or non-aromatic CRAM. For the terrigenous, aromatic-rich Congo River sample, almost all formulas with H/C below 1.2 were photo-labile, suggesting that a high proportion of these formulas were aromatic (Stubbins et al., 2010). For NADW photo-resistant formulas cover the full range of H/C values found in the initial sample, extending well below H/C 1.0 (Figure 3). Almost all (>95%) aromatic carbon, as indicated by CDOM, was photo-bleached out of the NADW samples (Stubbins et al., 2012). Therefore, the photo-resistance of NADW dissolved organics with ambiguous CRAM or lignin formulas, suggests a high proportion of these formulas are constituted to some degree by non-aromatic, non-photo-reactive CRAM-like moieties. Although hinting at a high number of CRAM-like molecular formula in NADW DOM, it must also be remembered that the current analysis considered only presence and absence and that each peak assigned a molecular formula

can be constituted of thousands of isomers. Therefore, the photo-resistance of a peak may derive from it containing both aromatic and non-aromatic isomers. During photo-degradation, the aromatic compounds would be degraded, but the FTICR-MS formula would persist along with the residual non-aromatic isomers. At AImod greater than 0.5 a formula has a very high chance of being aromatic. This is reflected in the high level of photo-lability of these formulas (Table 1; Figure 2B).

As for Congo River DOM, the photo-produced DOM was enriched in low AImod formulas, including aliphatic and peptide-like compounds (Table 1; Figure 2B and 3) (Stubbins et al., 2010). Aliphatic signatures in FTICR-MS analyses are highly bio-labile (Spencer et al., 2015), suggesting that the photo-production of aliphatics may play a role increasing the bio-lability of DOM upon photo-irradiation.

Conclusions

Photo-degradation alters the optical and chemical signature of DOM in surface waters. The photo-lability defined pools of DOM presented here offer novel insight into the molecular controls upon oceanic DOM photo-degradation. Of the trends reported, the loss of aromatic moieties is now well known and universally observed for DOM photo-degradation (Mopper et al., 2015). In the dark of the deep ocean, DOC is near-refractory and has accumulated over millennia to constitute a major pool of global carbon (Hansell, 2013). Aromatic molecules accumulate in the abyss, as indicated by increases in CDOM absorbance in the deep waters of the Pacific and Indian Oceans (Jørgensen et al., 2011; Nelson and Siegel, 2013) and DBC molecules (condensed aromatics) are evenly distributed in deep waters, but depleted due to photo-bleaching in surface waters (Dittmar and Paeng, 2009; Stubbins et al., 2012). Similarly, the oldest, or at least most radiocarbon depleted, fraction of oceanic DOM (30,000 years) is

elevated in the deep ocean (24 μM) relative to surface waters (9 μM) (Follett et al., 2014). Thus, one potential mechanism for the removal of ancient, deep ocean DOM is via photo-degradation when upwelling ventilates deep waters to the surface. If this is the case, then the magnitude of ancient, refractory carbon stored in the oceans may be constrained by the percentage of DOC that is photo-labile and the frequency with which it is cycled through sunlit waters. The presence of ancient DOC in surface waters indicates that photo-chemical stripping of ancient DOC from ventilated deep waters is not 100% efficient. Less than 100% removal in surface water could be explained by limitations in the efficiency of photo-chemical DOC mineralization and the length of time that an average DOC molecule will spend in sunlit waters. Alternatively, components of deep ocean, refractory DOC may be intrinsically resistant to photo-degradation.

For this reason, the photo-resistance of highly unsaturated formulas is of note (Table 1). The photo-refractory nature of these formulas is more apparent than for freshwater samples, such as the Congo River DOM that was irradiated under the same conditions (Stubbins et al., 2010), suggesting a higher percentage of these formulas or the isomers constituting these formulas within NADW DOM are non-aromatic, possibly CRAM-like compounds. Whatever their exact structure, the pool of highly unsaturated organics certainly contained photo-resistant isomers (Figure 3; Table 1). These compounds' photo-resistant nature would allow them to cycle through the surface ocean indefinitely. Compounds with the same molecular fingerprint are also found throughout the deep ocean and are suggested to comprise a major fraction of deep ocean refractory DOC (Hertkorn et al., 2006). Thus, these compounds may represent a pool of bio- and photo-resistant DOC that can accumulate almost indefinitely in the global ocean.

The near-complete photo-lability of CHONS compounds (97% of formulas disappeared) is a novel finding. Simple, non-aromatic CHONS compounds such as cysteine absorb UV light

(Johnson, 1971) and are photo-reactive in pure solution (Stubbins and Dittmar, 2015), suggesting that CHONS compounds may be routinely overlooked as a photo-reactive, light absorbing constituent of DOM. The full lists of photo-resistant, photo-labile, and photo-produced FTICR-MS assigned molecular formulas are presented in the Supplemental Information and are archived in the Pangaea database (Stubbins and Dittmar, 2015). This data will assist in identifying the molecular signatures of DOM photo-degradation and assigning photo-lability to molecular formulas within natural water samples.

Acknowledgements

We would like to thank the crew of the RV Atlantic Explorer and the Bermuda Institute of Ocean Sciences for assistance with sample collection. This work was supported by the National Science Foundation (OCE-1234704; OCE-1356890) and a Fellowship from the Hanse Institute for Advanced Studies (HWK, Delmenhorst, Germany) granted to Aron Stubbins.

Figure captions

Figure 1. Ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra for North Atlantic Deep Water before and after 28-day irradiation under a solar simulator. A) Initial sample, pre-irradiation. B) Post-irradiation.

Figure 2. A) Venn diagram representing molecular formulas identified in pre- and post-irradiation North Atlantic Deep Water. Area of overlap in black indicates photo-resistant molecular formulas present before and after irradiation. The blue area indicates photo-labile formulas present only pre-irradiation. The red area indicates photo-produced formulas, which were present only post-irradiation. B) Bar chart grouping the percentage of photo-resistant, photo-labile, and photo-produced molecular formulas by modified aromaticity index.

Figure 3. van Krevelen diagram where color indicates molecular mass. A) Photo-labile molecular formulas. B) Photo-produced molecular formulas. C) Photo-resistant molecular formulas.

Table 1. Molecular signatures in the pre- and post-irradiation NADW samples, as well as in the three photo-lability defined DOM classes (photo-resistant, photo-labile, photo-product). Percentages represent the distribution of each class of molecules among the different lability-defined classes of DOM.

	Pre-irradiation	Post-irradiation	Photo-resistant	Photo-labile	Photo-produced
All Formulas	3024	2402	2296 (73%)	728 (23%)	106 (3%)
Condensed Aromatics	2	0	0 (0%)	2 (100%)	0 (0%)
Aromatics	164	32	29 (17%)	135 (81%)	3 (2%)
Highly Unsaturated	2537	2109	2054 (79%)	483 (19%)	55 (2%)
Aliphatics	283	227	198 (63%)	85 (27%)	29 (9%)
Peptides	38	34	15 (26%)	23 (40%)	19 (33%)
CHO	1365	1237	1175 (82%)	190 (13%)	62 (4%)
With N	1308	992	949 (70%)	359 (27%)	43 (3%)
With S	346	137	135 (39%)	211 (61%)	2 (1%)
With P	78	40	40 (51%)	38 (49%)	0 (0%)
Average Mass	457	461	461	445	480
Average AI	0.25	0.24	0.24	0.27	0.13
Average H/C	1.26	1.27	1.26	1.25	1.44
Average O/C	0.47	0.48	0.48	0.43	0.51

Table 2. Relative contributions of each molecular class in the pre- and post-irradiation NADW samples, as well as in the three photo-lability defined DOM classes (photo-resistant, photo-labile, photo-product).

	Initial	Post	Photo-resistant	Photo-labile	Photo-product
Condensed Aromatics	0.1%	0.0%	0.0%	0.3%	0.0%
Aromatics	5.4%	1.3%	1.3%	18.5%	2.8%
Highly Unsaturated	83.9%	87.8%	89.5%	66.3%	51.9%
Aliphatics	9.4%	9.5%	8.6%	11.7%	27.4%
Peptides	1.3%	1.4%	0.7%	3.2%	17.9%
CHO	45.1%	51.5%	51.2%	26.1%	58.5%
With N	43.3%	41.3%	41.3%	49.3%	40.6%
With S	11.4%	5.7%	5.9%	29.0%	1.9%
With P	2.6%	1.7%	1.7%	5.2%	0.0%

References

- Bauer, J.E., 2002. Carbon isotopic composition of DOM. *Biogeochemistry of Marine Dissolved Organic Matter*: 405-453.
- Cherrier, J., Bauer, J.E., Druffel, E.R., Coffin, R.B. and Chanton, J.P., 1999. Radiocarbon in marine bacteria: Evidence for the ages of assimilated carbon. *Limnology and Oceanography*, 44(3): 730-736.
- Dittmar, T. and Paeng, J., 2009. A heat-induced molecular signature in marine dissolved organic matter. *Nature Geoscience*, 2(3): 175-179.
- Dittmar, T. and Stubbins, A., 2014. *Dissolved Organic Matter in Aquatic Systems, Treatise on Geochemistry: Second Edition*, pp. 125-156.
- Flerus, R. et al., 2012. A molecular perspective on the ageing of marine dissolved organic matter. *Biogeosciences*, 9(6): 1935-1955.
- Follett, C.L., Repeta, D.J., Rothman, D.H., Xu, L. and Santinelli, C., 2014. Hidden cycle of dissolved organic carbon in the deep ocean. *Proceedings of the National Academy of Sciences*, 111(47): 16706-16711.
- Gonsior, M. et al., 2009. Photochemically induced changes in dissolved organic matter identified by ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry. *Environmental science & technology*, 43(3): 698-703.
- Hansell, D.A., 2013. Recalcitrant dissolved organic carbon fractions. *Annual Review of Marine Science*, 5: 421-445.
- Hansell, D.A. and Carlson, C.A., 1998. Deep-ocean gradients in the concentration of dissolved organic carbon. *Nature*, 395(6699): 263-266.
- Helms, J.R. et al., 2014. Loss of optical and molecular indicators of terrigenous dissolved organic matter during long-term photobleaching. *Aquatic Sciences*, 76(3): 353-373.
- Helms, J.R. et al., 2013. Photochemical bleaching of oceanic dissolved organic matter and its effect on absorption spectral slope and fluorescence. *Marine Chemistry*, 155: 81-91.
- Helms, J.R. et al., 2008. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnology and Oceanography*, 53(3): 955-969.
- Hernes, P.J. and Benner, R., 2006. Terrigenous organic matter sources and reactivity in the North Atlantic Ocean and a comparison to the Arctic and Pacific oceans. *Marine Chemistry*, 100(1-2): 66-79.
- Hertkorn, N. et al., 2006. Characterization of a major refractory component of marine dissolved organic matter. *Geochimica et Cosmochimica Acta*, 70(12): 2990-3010.
- Hertkorn, N., Harir, M., Koch, B.P., 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(3): 1583-1624.
- Herzprung, P., Hertkorn, N., Friese, K. and Schmitt-Kopplin, P., 2010. Photochemical degradation of natural organic sulfur compounds (CHOS) from iron-rich mine pit lake pore waters—an initial understanding from evaluation of single-elemental formulae using ultra-high-resolution mass spectrometry. *Rapid Communications in Mass Spectrometry*, 24(19): 2909-2924.
- Johnson, E., 1971. *UV atlas of organic compounds*, 5.
- Jørgensen, L. et al., 2011. Global trends in the fluorescence characteristics and distribution of marine dissolved organic matter. *Marine Chemistry*, 126(1): 139-148.
- Koch, B.P. and Dittmar, T., 2006. From mass to structure: An aromaticity index for high-resolution mass data of natural organic matter. *Rapid Communications in Mass Spectrometry*, 20(5): 926-932.

- Koch, B.P., Dittmar, T., Witt, M. and Kattner, G., 2007. Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter. *Analytical Chemistry*, 79(4): 1758-1763.
- Kujawinski, E.B., Del Vecchio, R., Blough, N.V., Klein, G.C. and Marshall, A.G., 2004. Probing molecular-level transformations of dissolved organic matter: Insights on photochemical degradation and protozoan modification of DOM from electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Marine Chemistry*, 92(1-4 SPEC. ISS.): 23-37.
- Kujawinski, E.B. et al., 2002. The application of electrospray ionization mass spectrometry (ESI MS) to the structural characterization of natural organic matter. *Organic Geochemistry*, 33(3): 171-180.
- Medeiros, P. et al., 2015. Dissolved organic matter composition and photochemical transformations in the northern North Pacific Ocean. *Geophysical Research Letters*.
- Meyers-Schulte, K.J. and Hedges, J.I., 1986. Molecular evidence for a terrestrial component of organic matter dissolved in ocean water. *Nature*, 321(6065): 61-63.
- Mopper, K., Kieber, D.J. and Stubbins, A., 2015. Marine photochemistry: Processes and impacts. In: D.A. Hansell, Carlson, C. A. (Editor), *Biogeochemistry of Marine Dissolved Organic Matter*. Elsevier, pp. 389-450.
- Mopper, K., Stubbins, A., Ritchie, J.D., Bialk, H.M. and Hatcher, P.G., 2007. Advanced instrumental approaches for characterization of marine dissolved organic matter: Extraction techniques, mass spectrometry, and nuclear magnetic resonance spectroscopy. *Chemical Reviews*, 107(2): 419-442.
- Nelson, N.B. and Siegel, D.A., 2013. The global distribution and dynamics of chromophoric dissolved organic matter. *Annual Review of Marine Science*, 5: 447-476.
- Osburn, C.L., Morris, D.P., Thorn, K.A. and Moeller, R.E., 2001. Chemical and optical changes in freshwater dissolved organic matter exposed to solar radiation. *Biogeochemistry*, 54(3): 251-278.
- Powers, L.C. and Miller, W.L., 2015. Photochemical production of CO and CO₂ in the Northern Gulf of Mexico: Estimates and challenges for quantifying the impact of photochemistry on carbon cycles. *Marine Chemistry*, 171: 21-35.
- Singer, G.A. et al., 2012. Biogeochemically diverse organic matter in Alpine glaciers and its downstream fate. *Nature Geoscience*, 5(10): 710-714.
- Spencer, R.G.M. et al., 2014. Source and biolability of ancient dissolved organic matter in glacier and lake ecosystems on the Tibetan Plateau. *Geochimica et Cosmochimica Acta*, 142(0): 64-74.
- Spencer, R.G.M. et al., 2015. Detecting the signature of permafrost thaw in Arctic rivers. *Geophysical Research Letters*.
- Spencer, R.G.M. et al., 2009. Photochemical degradation of dissolved organic matter and dissolved lignin phenols from the Congo River. *Journal of Geophysical Research: Biogeosciences*, 114(3).
- Stenson, A.C., Landing, W.M., Marshall, A.G. and Cooper, W.T., 2002. Ionization and fragmentation of humic substances in electrospray ionization Fourier transform-ion cyclotron resonance mass spectrometry. *Analytical Chemistry*, 74(17): 4397-4409.
- Stubbins, A. and Dittmar, T., 2015. High resolution dataset accompanying "Illuminating the deep: Molecular signatures of photochemical alteration of dissolved organic matter from North Atlantic Deep Water". In: U.o.G. Skidaway Institute of Oceanography (Editor). *Pangaea*.
- Stubbins, A. et al., 2008. Relating carbon monoxide photoproduction to dissolved organic matter functionality. *Environmental science & technology*, 42(9): 3271-3276.
- Stubbins, A. et al., 2014. What's in an EEM? Molecular Signatures Associated with Dissolved Organic Fluorescence in Boreal Canada. *Environmental Science & Technology*, 48(18): 10598-10606.
- Stubbins, A., Niggemann, J. and Dittmar, T., 2012. Photo-lability of deep ocean dissolved black carbon. *Biogeosciences*, 9(5): 1661-1670.

- 606 Stubbins, A. et al., 2010. Illuminated darkness: Molecular signatures of Congo River dissolved organic
607 matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry.
608 *Limnology and Oceanography*, 55(4): 1467-1477.
- 609 Stubbins, A. et al., 2006. Open-ocean carbon monoxide photoproduction. *Deep-Sea Research Part II:*
610 *Topical Studies in Oceanography*, 53(14-16): 1695-1705.
- 611 Walker, B., Beaupre, S., Guilderson, T., Druffel, E. and McCarthy, M., 2011. Large-volume ultrafiltration
612 for the study of radiocarbon signatures and size vs. age relationships in marine dissolved organic
613 matter. *Geochimica et Cosmochimica Acta*, 75(18): 5187-5202.
- 614 Weishaar, J.L. et al., 2003. Evaluation of specific ultraviolet absorbance as an indicator of the chemical
615 composition and reactivity of dissolved organic carbon. *Environmental Science and Technology*,
616 37(20): 4702-4708.

617

618 Figure 1.

619

620

621

622

623

624

625

Figure 2

633 Figure 3

634

635

636

637