Carl von Ossietzky Universität Oldenburg
Fakultät für Mathematik und Naturwissenschaften

From freshly produced compounds to refractory molecules – tracing sources and fate of dissolved organic matter in the ocean

Von der Fakultät für Mathematik und Naturwissenschaften der
Carl von Ossietzky Universität Oldenburg zur
Erlangung des Grades und Titels eines Doktors der
Naturwisssenschaften
- Dr. rer. nat. –
angenommene Dissertation von Frau

Helena Osterholz
geboren am 26.12.1984 in Paderborn

1. Gutachter	Prof. Dr. Thorsten Dittmar
2. Gutachter	Prof. Dr. Meinhard Simon
3. Gutachterin	Dr. Katharina Pahnke
Tag der Disputation	03. Juli 2014

Table of Contents

Table of Contents I

Thesis abstract III

Zusammenfassung V

List of Abbreviations VIII

1 General Introduction 9
1.1 Introduction 9
1.1.1 What is dissolved organic matter? 9
1.1.2 Sources and sinks of DOM 16
1.1.3 Linking DOM and microorganisms 20
1.1.4 Current issues in DOM research 24
1.2 Objectives and outline of this thesis 27
1.3 Contributions to publications 28

2 Molecular evidence for rapid dissolved organic matter turnover in Arctic fjords 31
2.1 Highlights 32
2.2 Abstract 32
2.3 Introduction 33
2.4 Materials & Methods 35
2.5 Results 41
2.6 Discussion 47
2.7 Contributors 55
2.8 Acknowledgements 55
2.9 References 55

3 Do marine microorganisms really produce refractory dissolved organic matter? 61
3.1 Abstract 62
3.2 Results & Discussion 62
3.3 Methods summary 71
3.4 Author contributions 73
3.5 Acknowledgements 73
3.6 References 73
S1 Supplementary Methods 77
S2 Supplementary Results 87

S3 Supplementary References....91

4 Terrigenous input and microbial processing - driving forces of dissolved organic matter composition in the North Sea....95
4.1 Abstract....96
4.2. Introduction....96
4.3 Materials & Methods....100
4.4 Results & Discussion....104
4.5 Acknowledgements....119
4.6 References....119
4.7 Supplementary Results....126

5 Concluding remarks and perspectives....129

6 References....135

Curriculum Vitae....149

Acknowledgements....151

Eidesstattliche Erklärung....153

Thesis abstract

Dissolved organic matter (DOM) constitutes the biggest organic carbon pool in the ocean, fueling the marine carbon cycle. Most of the DOM is generated by autotrophic primary production and subsequently assimilated or transformed to sustain heterotrophic organisms. A minor part escapes degradation and enters the refractory DOM pool, where it remains unaltered for thousands of years. This thesis includes two field studies and one laboratory incubation experiment focusing on the production and turnover of freshly produced DOM by microorganisms. High resolution methods were applied to elucidate the DOM molecular composition and the microbial community structure, and the data was analyzed via multivariate statistics to provide novel insights into the relationship between the two key players.

DOM turnover at high latitudes was studied in fjords of Svalbard, Norway. The sampling campaign took place approximately 2 months after the decline of the phytoplankton blooms which usually occur from April to June in the archipelago. Denaturing gradient gel electrophoresis (DGGE) fingerprinting revealed differences in microbial community composition with location and water depth. The molecular DOM fingerprint obtained via ultrahigh-resolution mass spectrometry (Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, FT ICR MS), however, did not significantly differ between surface water, which had recently experienced primary production, and the more saline, N-rich bottom waters. Consequently, the freshly produced DOM must be degraded within the short, but despite the cold conditions, highly productive warm season by the resident microbial community. The refractory DOM background found at most oceanic locations worldwide persists in the fjord water masses.

In order to test whether this universal signature of deep sea DOM which is considered a result of decadal to millennial transformation can be produced solely by microbially mediated reworking of DOM on timescales of weeks to months, a long-term batch incubation study was carried out. The production and degradation of the freshly produced organic matter by the microbial community was followed over

the course of almost three years and bulk as well as molecular DOM characteristics were monitored. After only ~1 month of incubation, the molecular DOM composition was highly diverse and on the level of presence/absence of assigned molecular formulae exhibited high coherence with deep ocean refractory DOM. In-depth analysis via ultrahigh resolution mass spectrometry taking into account the characteristic intensity distribution of deep ocean DOM revealed that, independent from the community composition of the producers, only up to 3 μmol C L^{-1}, or 0.2-0.4% of the net community production, can be considered truly refractory marine DOM. This estimate is in the same order of magnitude as the rates of global refractory DOM production, suggesting that biological processes alone produce sufficient amounts of biologically resistant DOM to sustain the marine refractory DOM pool.

DOM availability may restrict the abundance of certain bacterial taxa, while microorganisms shape the composition of the DOM pool through selective uptake or production. Insight into the relationship of molecular DOM composition and microbial community structure - examined via pyrosequencing of 16S rRNA amplicons of environmental RNA and DNA - was gained through a third study: Within the pronounced salinity gradient of the North Sea, the DOM composition and the total microbial community (DNA based) composition were strongly linked and driven by freshwater input at the southern coastline. The active microbial community (RNA based) showed a weaker link to DOM and was less impacted by salinity-driven changes. The proposed statistical approach represents the first combination of results from high resolution mass spectrometry and a next generation sequencing approach, providing a good basis for the exploration of similar, even larger datasets.

Together, the results obtained from field and laboratory studies indicate a rapid turnover of freshly produced DOM by the respective resident microbial community that might ultimately result in the universal background signature observed in oceanic DOM.

Zusammenfassung

Gelöstes organisches Material (engl. dissolved organic matter, DOM) als größtes organisches Kohlenstoffreservoir im Meer ist von entscheidender Bedeutung für den marinen Kohlenstoffkreislauf. Der vorwiegende Teil des DOM wird von autotrophen Primärproduzenten erzeugt und anschließend zum Aufbau von Biomasse genutzt oder dient heterotrophen Organismen als Nahrungsgrundlage. Ein kleiner Anteil entgeht jedoch dem Abbau und bleibt als refraktäres DOM über tausende von Jahren erhalten. In dieser Arbeit werden zwei Feldstudien sowie ein im Labor durchgeführtes Experiment beschrieben, deren Hauptaugenmerk auf der Produktion und dem anschließenden Abbau des neu produzierten DOM durch Mikroorganismen liegt. Anhand hochauflösender Methoden wurden die molekulare Zusammensetzung des DOM und die Struktur der entsprechenden mikrobiellen Gemeinschaft untersucht und die Ergebnisse anschließend mit multivariater Statistik ausgewertet um neue Einblicke in die wechselseitigen Beziehungen der beiden wesentlichen Akteure des Kohlenstoffkreislaufes zu erlangen.

Die Umsetzung von DOM in Gewässern hoher Breitengrade wurde in den Fjorden von Svalbard, Norwegen, untersucht. Die Probennahme fand etwa zwei Monate nach den Phytoplanktonblüten statt, die in den Gewässern um die Inselgruppe zwischen April und Juni ihr Maximum erreichen. Die Bakteriengemeinschaft, beschrieben anhand der denaturierenden Gradienten-Gelelektrophorese (DGGE) zeigte je nach Standort und Wassertiefe Unterschiede in ihrer Struktur. Ultrahochauflösende Massenspektrometrie (Fouriertransformation Ionenzyklotron-resonanz Massenspektrometrie, FT ICR MS) ließ keinerlei Unterschiede in der molekularen Zusammensetzung des DOM zwischen dem Oberflächenwasser, in dem unlängst die Phytoplanktonblüten herrschten, und dem stickstoffreicheren Tiefenwasser mit höherem Salzgehalt erkennen. Folglich muss das unlängst neu produzierte DOM im kurzen, trotz niedriger Temperaturen aber hochproduktivem, Sommer von der Bakteriengemeinschaft umgesetzt worden sein. Das Signal des

refraktären DOM, welches in allen Ozeanen der Welt zu finden ist, besteht auch in den Wassermassen der Fjorde.

Diese universale Signatur des Tiefsee-DOM kann durch Transformationsprozesse im Laufe tausender Jahre herausgebildet werden. Ob auch mikrobieller Umsatz innerhalb weniger Monate oder Jahre DOM einer ähnlichen Zusammensetzung hervorbringt, wurde anhand einer Langzeitinkubationsstudie untersucht. Die Produktion und der anschließende Abbau des organischen Materials durch eine gemischte Bakterien- und Phytoplanktongemeinschaft wurde über einen Zeitraum von fast drei Jahren beobachtet und die Zusammensetzung des DOM sowohl anhand von Summenparametern als auch auf molekularer Ebene verfolgt. Nach nur einem Monat wies das DOM eine hohe molekulare Diversität auf und ein hoher Anteil der zugeordneten Summenformeln stimmte mit denen von refraktärem Tiefsee-DOM überein. Eine detailliertere Analyse bei der auch die charakteristische Intensitätsverteilung des refraktären DOM in den ultrahochaufgelösten Massenspektren herangezogen wurde zeigte jedoch, dass maximal 3 µmol C L^{-1}, oder 0.2-0.4% der Nettoproduktion der Gemeinschaft, als wirklich refraktäres DOM angesehen werden können. Diese Abschätzung liegt in der gleichen Größenordnung wie die Raten der globalen Produktion von refraktärem DOM und zeigt so, dass biologische Prozesse allein ausreichen um die Größe des refraktären, organischen Kohlenstoffreservoirs im Ozean zu erhalten.

Die Verfügbarkeit des DOM kann einerseits das Vorkommen bestimmter Arten von Bakterien einschränken, andererseits können Mikroorganismen die Zusammensetzung des DOM durch selektive Aufnahme oder Produktion bestimmter Komponenten verändern. Ein Einblick in die Zusammenhänge der molekularen Zusammensetzung des DOM und der Struktur der mikrobiellen Gemeinschaft - ermittelt anhand von Pyrosequenzierung der 16S rRNA Amplifikate aus RNA- und DNA-Extrakten – wurde durch eine dritte Studie ermöglicht. Entlang des ausgeprägten Salzgradienten der Nordsee zeigte sich eine hochsignifikante Korrelation der molekularen Zusammensetzung des DOM und der Gesamtbakteriengemeinschaft (DNA-basiert), die offensichtlich vom Süßwassereintrag an den Küsten beeinflusst wurde. Die aktive Bakteriengemeinschaft (RNA-basiert) war

weniger deutlich mit der molekularen DOM Zusammensetzung verknüpft und auch der Einfluss des Salzgradienten war wesentlich geringer. Die herausgearbeitete statistische Herangehensweise kombiniert erstmals Ergebnisse hochauflösender Massenspektrometrie mit denen des „next generation sequencing“ und stellt damit eine solide Basis für die weitere Auswertung dieser Art von Datensätzen dar.

Die Ergebnisse der Arbeit lassen insgesamt darauf schließen, dass Mikroorganismen eine Schlüsselrolle in der schnellen Umsetzung von frisch produziertem DOM spielen, die letztendlich auch die universale Signatur des refraktären DOM-Hintergrunds der Ozeane herbeiführt.

List of Abbreviations

16S rRNA	16S ribosomal ribonucleic acid
BPCA	Benzenepolycarboxylic acid
CTD	Sonde to determine conductivity, temperature, and depth of the ocean
DBC	Dissolved black carbon
DBE	Double bond equivalent
DGGE	Denaturing gradient gel electrophoresis
DNA	Desoxyribonucleic acid
DOC	Dissolved organic carbon
DOM	Dissolved organic matter
DON	Dissolved organic nitrogen
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
FDNS	Free dissolved neutral sugars
FT ICR MS	Fourier transform ion cyclotron resonance mass spectrometry
HMW	High molecular weight
HPLC	High performance liquid chromatography
ICBM	Institute for Chemistry and Biology of the Marine Environment
LMW	low molecular weight
m/z	Mass to charge ratio
MC	Microbial community
MF	Molecular formulae
MUC	Molecularly uncharacterized component
NMR	Nuclear magnetic resonance
NPDW	North Pacific Deep Water
NSW	North Sea Water
PC	Polycarbonate
PCA	Principal component analysis
PCoA	Principal coordinate analysis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
S/N	Signal to noise ratio
SPE	Solid phase extraction
TDN	Total dissolved nitrogen
TEP	Transparent exopolymer particles
THDAA	Total hydrolysable dissolved amino acids

1 General Introduction

1.1 Introduction

The first section provides an overview of the global dissolved organic matter (DOM) distribution and summarizes the current state of knowledge in DOM composition. Furthermore, the sources and sinks will be discussed to describe the role of DOM in the Earth's biogeochemical cycles. The following section begins with a short introduction to molecular microbial community analysis before the linkage between DOM and microbial community composition is introduced. The last section describes topics presently discussed in the field of DOM research that were touched upon in at least one of the studies included in this thesis.

1.1.1 What is dissolved organic matter?

Carbon is often termed the building block of life. This seems justified when considering that it comprises about half of the total dry mass of living things on Earth. But carbon is also abundant in the non-living reservoirs in the Earth's crust, soils and the atmosphere. The global carbon cycle describes the path of carbon-containing compounds through the atmosphere, continents and oceans. Chemical transformations involving plants and microorganisms play an important role in this cycle, linking abiotic and biotic carbon pools on different timescales.

Embedded in this frame of the global carbon cycle, and impossibly considered separately, the dissolved organic matter (DOM) in the oceans plays a vital role as one of the largest *active* carbon stocks. With 662 Pg C, it comprises a similar amount of carbon as is bound in CO_2 in the atmosphere (Hansell et al. 2009; Hedges 1992), but its composition and role in biogeochemical cycles are far less understood. The distribution of dissolved organic carbon (DOC) as a quantitative measure of DOM throughout the world's oceans then again has been studied in increasing spatial and temporal resolution since DOM has first been recognized in the literature

by the German scientist August Pütter in 1907. Today researchers can paint a relatively clear picture of the DOC production in the euphotic zone that leads to higher concentrations in the surface waters of the ocean and the DOC transport and subsequent degradation/dilution along the path of thermohaline circulation into the deep ocean at North Atlantic or Southern Ocean downwelling sites (Fig. 1, Hansell 2013).

Understanding the molecular composition of DOM is crucial in elucidating the driving forces behind DOC concentrations and the storage/fate of DOM. Categorization of DOM compounds into classes based on size, origin, reactivity, or chemical composition, simplifies communication and also aids in the understanding of the role of each class in the environment.

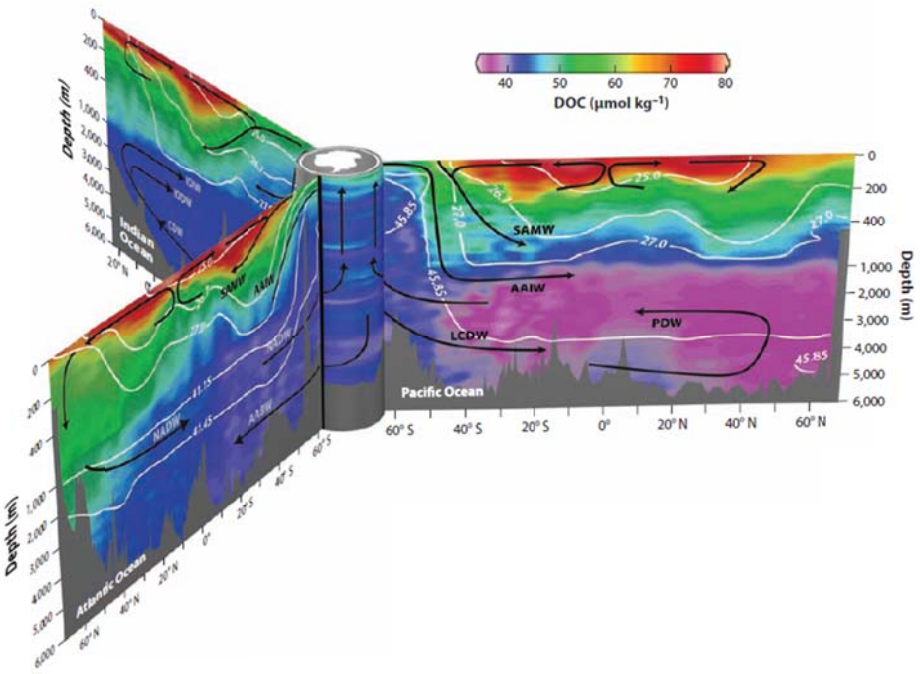

Fig. 1. DOC concentrations (color scale), major currents (black arrows) and isopycnal surfaces (white lines) in the global oceans (from Hansell 2013). DOC concentrations at the surface are high due to primary production and release of DOC, which is then degraded and diluted as it is transported along the major ocean currents.

The size continuum of DOM. The classification by size offers, at first view, the easiest way to separate DOM fractions. The most widespread definition of DOM

includes organic matter that passes through a 0.1 to 1 µm filter. This operational definition isolates particulate organic matter (POM) and transparent exopolymer particles (TEP), formed by the spontaneous polymerization of acidic polysaccharides (Alldredge et al. 1993), from free polymers and DOM that can further be divided into low molecular weight (LMW) and high molecular weight DOM (HMW, Fig. 2). 60 to 80% of the bulk open ocean DOM are LMW-DOM that is usually smaller than 1 kDa (Ogawa and Tanoue 2003). LMW-DOM may consist of the biologically labile monomer carbohydrates and amino acids, but also makes up the largest part of highly refractory DOM in the deep ocean with lifetimes of ~16,000 years (Hansell 2013). HMW-DOM and colloidal material span the range between truly dissolved and particulate organic matter. 20 to 35% of the bulk DOM is larger than 1 kDa while 2 to 7% exceed the size of 10 kDa and are thus assigned to this fraction. A more appropriate attribution of DOM into a size continuum has been proposed (Amon and Benner 1996), owing to the fact that e.g. the marine gelphase can span all size classes (Verdugo et al. 2004) and the finding that functional features are poorly represented by the size alone.

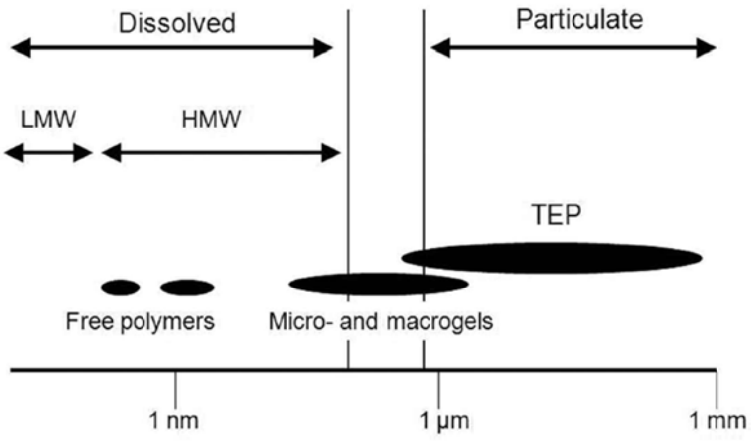

Fig. 2. The size continuum of marine organic matter (from Hunt et al. 2010). Dissolved organic matter is often operationally defined as <0.7 µm and can be separated into low molecular weight (LMW) and high molecular weight (HMW) fractions, also including free polymers. Transparent exopolymeric particles (TEP) are included in the particulate phase, while micro- and macrogels span the intermediate size fraction.

Reactivity fractions of DOM. The partitioning of DOM into fractions with discrete reactivity (Hansell 2013) allows more consideration of the inherent properties of DOM. On the downside, the groups are less easily distinguished by analytical methods as the classification relies mostly on removal times and relative contribution to bulk DOC. The reactivity of organic molecules in the ocean ranges from labile compounds that are turned over within minutes to days, to the most refractory compounds that persist for millennia (Fig. 3). Small amino acids, sugars, and short-chain fatty acids are easily taken up by bacteria and mostly escape analytical detection due to their short residence times. Semi-labile DOM is mostly present in surface waters for several months up to 1.5 years in the euphotic zone above a seasonal pycnocline. Semi-refractory DOM is stable for about 20 years and generally observed in the upper ocean layer (0-1000 m), but its production and removal processes are largely unknown. While the labile and semi-refractory fractions occur almost exclusively at water depths above 1000 m, refractory and ultra-refractory DOM are ubiquitously distributed in the ocean. With ~630 Pg C, the refractory DOM pool accounts for the largest fraction. The lifetimes of single compounds in this pool may exceed 16,000 years and thus surpass the circulation time of the ocean several times. The bulk age has been determined to vary between 4,000 to 6,000 years due to mixing of compounds of different ages (Bauer et al. 1992; Williams and Druffel 1987). Recently, the ultra-refractory DOM which persists for ~40,000 years was differentiated as the least reactive carbon fraction (Dittmar et al. 2012; Dittmar and Paeng 2009). Evidence exists that this kind of DOM is formed during the combustion of biomass on land (Jaffé et al. 2013) or may be formed during burial in deep ocean sediments, representing a link to the inactive carbon pools such as kerogen (Killops and Killops 1993), that undergo only small changes on geological timescales.

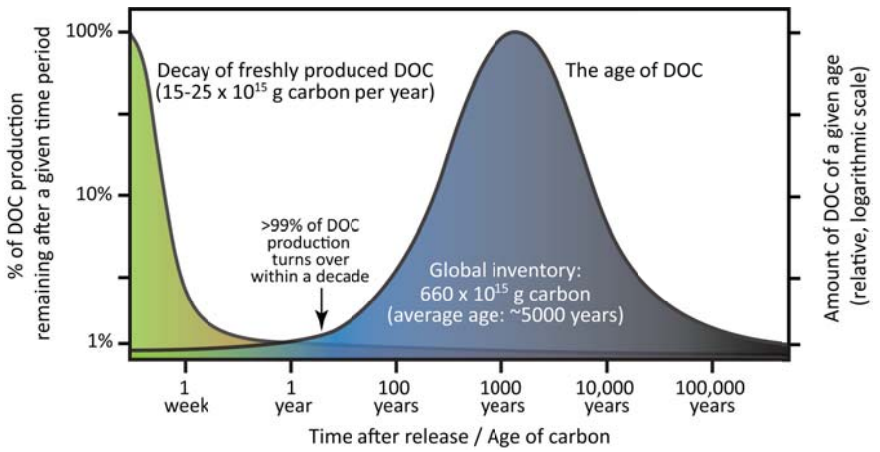

Fig. 3. Conceptual scheme illustrating the age distribution of marine DOC (blue) and the decay of freshly released DOC (green). Most DOC is quickly consumed, a small fraction of DOC accumulates in the ocean over thousands of years (from Dittmar 2014).

Compound classes in DOM. Ultimately, chemical characterization of DOM on molecular level is needed to understand its pathways through global biogeochemical cycles. Bulk chemical characterization, e.g. the C/N ratios of different fractions have shown that surface water is usually enriched in organically bound N which is then preferentially decomposed (Hopkinson et al. 1997; Loh and Bauer 2000; Sambrotto et al. 1993). In more detail, molecularly defined building blocks such as amino acids, carbohydrates and fatty acids, representing mostly the labile, bioavailable quota, can be quantified with chemical methods due to their known structural composition. This fraction rarely exceeds 30% of bulk DOM (Mccarthy et al. 1996; Skoog and Benner 1997), but provides valuable insight into the early stages of DOM cycling. On the basis of amino acids, for example, a degradation state index (Dauwe and Middelburg 1998) can be calculated using the relative contributions of several amino acids that vary in nitrogen content, size or functional groups and are thus more or less effectively used by microorganisms. Similarly to amino acids, combined neutral sugars provide information about the diagenetic state of DOM. The sugars are produced by phytoplankton at the sea surface where enough light is

available and then are readily used by heterotrophs (Amon and Benner 2003). So far, little is known about fatty acids in DOM, but it has been shown that they provide information about the source of the DOM since they are important constituents of bacterial membranes, especially in Gram negative species (Wakeham et al. 2003). The fact that specific DOM compounds can be traced back to their origin reveals their high biomarker potential (Mccallister et al. 2006). This is also true for lignin phenols: rivers do not only supply nutrients, but also terrigenous organic matter to the oceans. A special case is the Arctic Ocean, which, despite its small size, receives 10% of the global freshwater input (Benner et al. 2005). The lignin phenol concentration and composition, an unambiguous marker for vascular plant material, were traced by Hernes and Benner (2006) from the Arctic through the Atlantic into the Pacific and decreasing concentrations hint towards diagenesis and mixing of these compounds. In the North Atlantic, these compounds account for 1 to 2% of the bulk DOC concentration. Moreover, black carbon, a product of incomplete combustion, has for a long time only been recognized as an important component of soil particulate organic matter. Marine dissolved black carbon (dBC) has now been recognized to contribute to ultra-refractory DOM (Hansell 2013) and is thought to account for about 10% of the global riverine DOC flux (Jaffé et al. 2013).

However, the oceanic DOM pool also includes thousands of compounds that are of unknown structural composition and whose diversity exceeds the resolving power of established methods (Dittmar et al. 2007; Woods et al. 2011). The combination of desalting and concentration methods such as solid phase extraction (SPE, Dittmar et al. 2008), ultrafiltration (Benner et al. 1992) or reverse osmosis coupled with electrodialysis (RO/ED, Vetter et al. 2007) of seawater with ultrahigh resolution mass spectrometry, namely Fourier-transform ion cyclotron resonance mass spectrometry (FT ICR MS), has made it possible to address the composition of the "molecularly uncharacterized component" (MUC, Loh et al. 2006) in natural waters. Several compounds of different elemental composition can now be resolved per nominal mass and the high resolution and mass accuracy enable the assignment of molecular formulae to thousands of peaks per sample employing the mass defect (Fig. 4). DOM compounds consist mainly of combinations of the elements C, H, N,

O, P and S, which are commonly used for formula assignment (Koch et al. 2007). With only these few elements, we can learn a lot about the previously uncharacterized fraction of DOM. This technique has been used to elucidate the characteristics of terrigenous and marine DOM (Koch et al. 2005), distinguish DOM of different ages in oceanic water masses (Flerus et al. 2012) or trace the influence of photochemically induced reactions (Stubbins et al. 2010).

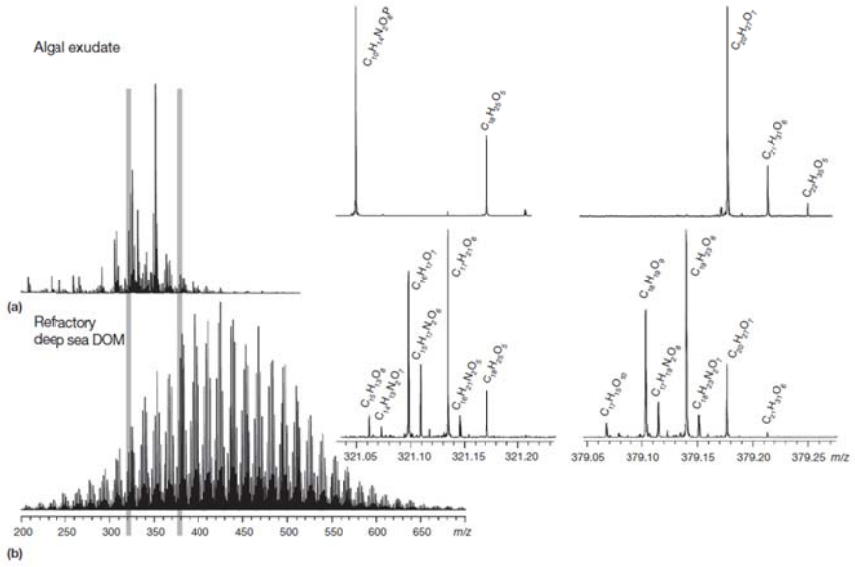

Fig. 4: Ultrahigh-resolution mass spectra of (a) labile DOM and (b) refractory DOM obtained on a 15 Tesla Bruker Solarix FT ICR MS. The whole mass spectra and two exemplary nominal masses (321 and 379) are shown. Refractory DOM is characterized by an enormous molecular diversity and very regular, recurring mass spacing patterns. Labile DOM is less diverse and characterized by chaotic abundance patterns of the individual molecules (from Dittmar and Stubbins 2014).

With all its advantages, the information obtained by FT ICR MS is restricted to molecular formulae. Only hydrogen and oxygen deficient DOM molecules can be identified to the structural level through this method (Dittmar and Paeng 2009), while all other compounds with assigned molecular formulae can in theory harbor thousands of structures each (Hertkorn et al. 2013). It is, however, possible to derive structural information via tandem FT ICR MS analyses. The fragmentation pattern of a nominal mass or even a single peak provides insight into the prevailing

functional groups (Liu et al. 2011; Witt et al. 2009). Nuclear magnetic resonance (NMR), in combination with FT ICR MS, is a very suitable addition to the in depth characterization of DOM since it can provide information about key structural components in a sample (Mopper et al. 2007). Using this combined approach, Hertkorn et al. 2013 found for example that surface Atlantic Ocean DOM contains fewer methyl esters than the deep waters and that carboxylic acids and ketones increase with depth.

The beneficial combination of all available chemical analytical methods, from bulk parameters over molecular composition to the point of structure determination has led and will continue to lead to an improved understanding of the marine DOM composition, which in turn is the base for the reconstruction of biogeochemical cycles.

1.1.2 Sources and sinks of DOM

DOM sources - autochthonous production. The dominant source of DOM in the ocean is photosynthesis in the euphotic zones (Baines and Pace 1991; Fogg 1983; Pomeroy 1974). Photoautotrophic organisms take up CO_2 and H_2O to build organic compounds using the energy from light. Organic carbon compounds derived from chemoautotrophic fixation in the deep ocean also contribute to the autochthonous DOM production, but are less explored and presumably amount to less than 1 Pg C yr^{-1} (Middelburg 2011). Overall, 45 to 50 Pg C are fixed by phytoplankton, which in its biomass only comprises roughly 1 Pg C, every year (Carr et al. 2006). Much of this organically bound carbon is needed for respiration, so that the net uptake of CO_2 comes out at only 2±1 Pg C yr^{-1}. According to Ducklow and Carlson (1992), about 50% of the carbon fixed by phytoplankton then reaches the DOM pool through various processes: Phytoplankton may use the carbon that is not needed to sustain their metabolism for cell growth or reproduction. Excess carbon is released by the cells at variable rates depending on community structure, environmental conditions or physiological state (Fogg 1983; Myklestad 1995; Wetz and Wheeler 2007). DOM is furthermore released through the grazing by zooplankton - sloppy feeding,

excretion and egestion release DOM in the environment during the process (Jumars et al. 1989; Nagata and Kirchman 1992). Bacterial or phytoplankton cell lysis due to viral infection and bacterially induced phytoplankton lysis are additional factors contributing to DOM production (Lonborg et al. 2013). POM and larger DOM molecules are effectively broken down by bacterial ectoenzymes and then released as DOM (Arnosti et al. 2011; Smith et al. 1995).

DOM sources - allochthonous input. Most of the allochthonous carbon reaches the ocean by riverine inflow, groundwater discharge or wet and dry atmospheric deposition.

DOM is released from particulate organic matter by leaching processes in the soil system or within the water column, which is accompanied by fractionation processes such as the sorption to mineral surfaces (Hedges et al. 1994). By the time the DOM reaches the ocean, it carries a distinct terrigenous signature including a high C/N ratio, a stable carbon isotopic composition with low $\delta^{13}C$ values and relatively high concentrations of lignin phenols as unambiguous tracers of vascular plant material (Lobbes et al. 2000). The world's rivers discharge about 0.4 Pg C yr^{-1} in dissolved and particulate form into the oceans (Schlesinger and Melack 2011). Roughly half of this is contributed in the form of DOM (0.25 Gt C yr^{-1}, Cauwet 2002; Hedges et al. 1997) Compared to the marine primary production, this input is small, but may provide an important carbon source for bacteria in the pelagic zones (Tranvik 1992). However, also large amounts of highly degraded, nitrogen-poor remains of terrestrial organisms (Hedges et al. 1997) and black carbon (Jaffé et al. 2013), representing recalcitrant DOM fractions enter the ocean via the rivers.

Generally, submarine groundwater discharge represents a diffuse flow and is as such not easily quantified. The current lack of knowledge is illustrated by a global estimate of the amount of water transported via these systems that ranges from 0.01 to 10% of the surface runoff (Taniguchi et al. 2002), spanning three orders of magnitude. One of the few studies tracing the contribution of DOC by the diffuse input was conducted by Kim et al. (2012), who were able to show that the

subterranean estuary of Hampyeaong Bay, Korea is indeed a net source of DOC, introducing 2 to 5×10^9 g C yr^{-1} into the bay.

Atmospheric dust deposition is mainly regarded as a source of minerals and trace metals to the ocean (Lawrence and Neff 2009). That organic carbon deposition may also be of ecological relevance has so far only been recognized for C-limited alpine catchments, where amount of carbon inputs via wet and dry deposition similar to the C contribution from microbial autotrophic production in barren soils have been reported (Mladenov et al. 2012). The analysis of a particular form of wet deposition, the hailstones, revealed the existence of diverse DOM with low fractions of easily biodegradable compounds such as carbohydrates or lipids (Šantl-Temkiv et al. 2013).

DOM sinks – the biological pump. The biological pump describes the ocean's mechanism of biologically driven carbon sequestration (Sigman and Haug 2003; Volk and Hoffert 1985). Atmospheric CO_2 is dissolved in the surface ocean and through deep water formation at high latitudes transported into the ocean interior (Fig. 5). Primary producers take up the reduced carbon in the photic zones for use in photosynthetic processes and divert it to organic matter at a ratio of C:N:P of 106:16:1, the Redfield ratio (Redfield 1934). The organisms and non-living particles may stay in the euphotic zone or slowly sink through the water column, 1 to 40% of them reaching the dark ocean (Herndl and Reinthaler 2013). On their way, they are slowly degraded by heterotrophic organisms or form larger aggregates, that way increasing their sinking velocity and escaping degradation. Once the organic matter arrives at the sea floor, most of the carbon will have been consumed by microorganisms, and remineralization will have released nutrients that can then again be used by primary producers. Only 1% of the global primary production will be buried in the sediments and remain there for thousands of years. Embedded within the biological pump, the microbial loop (Azam et al. 1983) describes the return of mainly phytoplankton-derived DOC into the food chain and through the uptake by bacteria and incorporation into biomass. The bacteria are in turn consumed by zooplankton, which serves as nutrition for higher trophic levels.

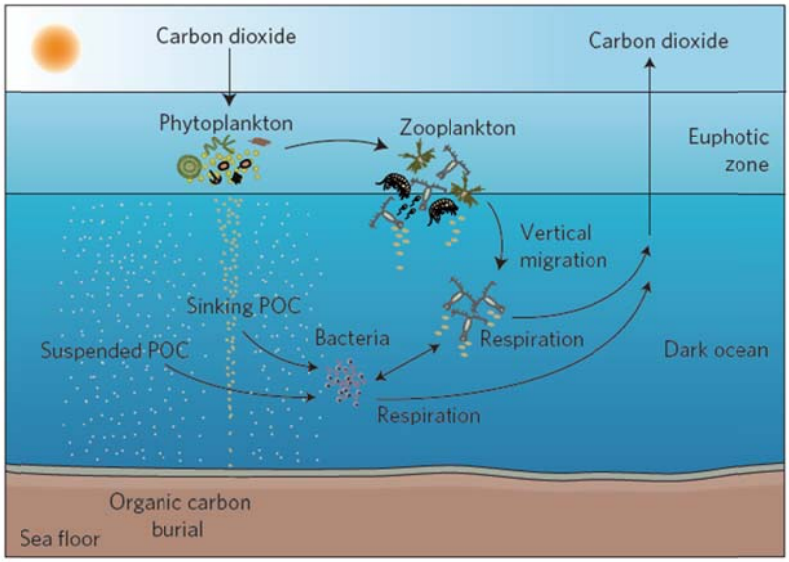

Fig. 5. The biological pump describes the microbially mediated flux of carbon from the atmosphere to the ocean interior. Phytoplankton in the euphotic zone fix CO_2 using solar energy. The POC is grazed on by herbivorous zooplankton or consumed by heterotrophic microbes. Remineralization of organic matter in the oceanic water column converts the organic carbon back to carbon dioxide (from Herndl and Reinthaler 2013)

DOM sinks – abiotic processes. Photochemical processes may lead to the direct mineralization of DOM to CO_2 (Mopper et al. 1991) and thus decrease the DOC concentration at the ocean's surface especially at low latitudes. UV excitation may also render the susceptibility of DOM to biological transformation. Different studies have shown an enhanced degradability by microbes after UV exposure (Kieber et al. 1990; Moran and Zepp 1997), the production of more refractory DOM (Naganuma et al. 1996) or no effect on DOM concentration at all (Thomas and Lara 1995). The different conclusions from these studies and experiments suggest that the influence of solar radiation on DOM is rather complex and may consist of several interconnected processes, such as the coupling of UV radiation and humic matter in the production of refractory DOM (Tranvik and Kokalj 1998). A second abiotic process leading to the removal of DOM is the sorption to either biogenic particles such as aggregates and diatom frustules or to mineral surfaces (Keil et al. 1994; Satterberg et al. 2003). These processes are probably of highest importance in areas

of extensive phytoplankton blooms close to the coast and in sediments, where 90% of the DOM are immediately scavenged by minerals (Keil et al. 1994). Adsorption to biogenic particles that then sink through the water column leads to rapid burial in the ocean where the aggregated DOM is then only slowly degraded. Adsorption to minerals physically protects DOM from bacterial uptake because it may enclose the DOM in pores smaller than 10 µm (Mayer 1994), too small to allow the functioning of hydrolytic enzymes.

1.1.3 Linking DOM and microorganisms

Culture-independent methods in molecular microbial community analysis. Microorganisms dominate the biosphere, totaling up to 4 to 6×10^{30} prokaryotic cells (Whitman et al. 1998). Hence, they represent key players in the recycling of elements and nutrients, while their genetic diversity is largely unexplored. Culture-based methods currently capture less than 1% of the community (Hugenholtz 2002). Over the last decades molecular, culture-independent methods that are based on the comparative analysis of rRNA gene signatures, especially of the highly conserved 16S rRNA gene, have advanced. In the following, only the denaturing gradient gel electrophoresis (DGGE) fingerprinting and next-generation sequencing 454 pyrosequencing approaches will be exemplified as they are most relevant to the work contained in this thesis.

Usually, DNA or RNA is extracted from an environmental sample and the gene region of interest is amplified via a polymerase chain reaction (PCR, Mullis and Faloona 1987). The PCR products can then directly be submitted to genetic fingerprinting methods such as DGGE. The DNA templates, amended with a 5'-GC clamp to prevent complete cleaving of the DNA strand, are electrophoresed on a polyacrylamide gel containing a linear gradient of a DNA denaturant. The nucleotide sequence of the DNA determines their melting behavior and the amplicons cleave and stop migration at different positions in the gel (Fig. 6, Muyzer et al. 1993). The gel is then stained, photographed and the fingerprints of the samples compared using computer-based clustering methods. Furthermore, single

bands may be excised, reamplified, and sequenced to obtain more detailed phylogenetic assignment. The advantages of this method are a fast overview over the diversity of several samples and the relatively low cost. On the downside, DGGE may not reflect the total diversity of a sample since only the abundant species are detected that make up more than 1 to 2% of the total community (Murray et al. 1996; Muyzer et al. 1993) or may overestimate the diversity of the microbial community due to sequence heterogeneities of multiple rRNA operons of one single strain (Nübel et al. 1996).

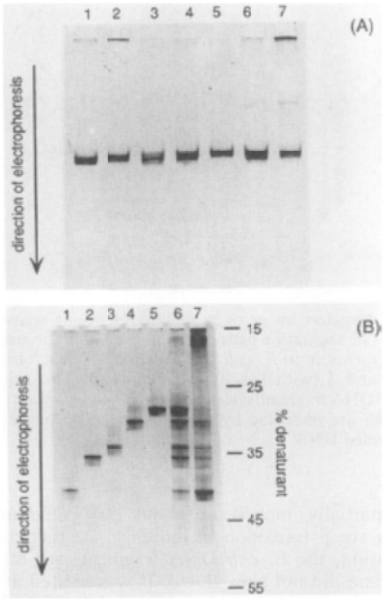

Fig. 6. Neutral polyacrylamide (A) and DGGE (B) analyses of 16S rDNA fragments of different eubacteria obtained after PCR amplification. Negative images of ethidium bromide-stained separation patterns of *D. saprovans* (lane 1), E. coli (lane 2), *M. chthonoplastes* (lane 3), *T. thioparus* (lane 4), *D. desulfricans* (lane 5), a mixture of these PCR products (lane 6), and a sample obtained after enzymatic amplification of a mixture of the bacterial genomic DNAs (lane 7, from Muyzer et al. 1993).

Likewise a PCR-based approach, 454 pyrosequencing offers a much more detailed insight into the diversity, but also the phylogenetic composition of a microbial community. The "sequencing by synthesis" principle has been further developed and improved since its discovery in 1988 (Ronaghi et al. 1998). It is based on the detection of the chemoluminescent release of pyrophosphate during DNA synthesis. The different nucleotides are added sequentially and the amount of light released is

proportional to the number of nucleotides that were incorporated (Fig. 7). The short sequence lengths generated by this particular next generation sequencing technique that may render unambiguous phylogenetic assignment via databases difficult (Siqueira et al. 2012), the comparably high cost and the extensive data processing after analysis constitute some of the drawbacks of this method. However, the high resolution enables the investigation of microbial community composition in unprecedented detail, e.g. of the largely understudied rare biosphere (Galand et al. 2009; Zhan et al. 2013).

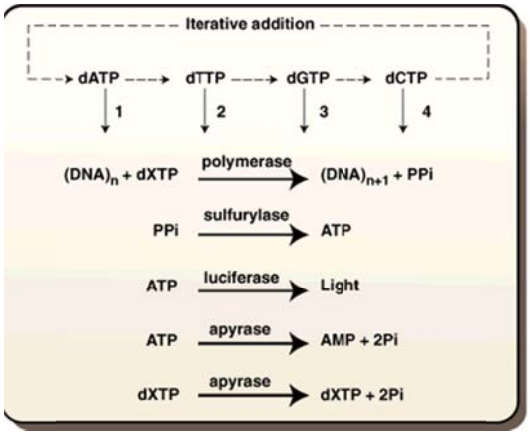

Fig. 7: Four nucleotides are added stepwise to the template hybridized to a primer. The PPi (pyrophosphate) released in the DNA polymerase-catalyzed reaction is detected by the ATP sulfurylase and luciferase in a coupled reaction. The added nucleotides are continuously degraded by a nucleotide-degrading enzyme. After the first added nucleotide has been degraded, the next nucleotide can be added. As this procedure is repeated, longer stretches of the template sequence are deduced. dXTP, one of the four nucleotides (from Ronaghi et al. 1998).

Linking microbial community structure to DOM composition – the current state of knowledge. Dissolved organic matter is the primary substrate that supports bacterial metabolism in the pelagic zones of the oceans. Through its high diversity, DOM supports the diversity of the marine microbial community, which in turn exerts influence on the major biogeochemical cycles, primary production and the marine food web. Researchers are just starting to decipher the complexity of both, molecular DOM composition and microbial community structures - a key requisite for the understanding of the numerous interactions. Technological advances such as

high throughput pyrosequencing and metatranscriptome analyses and ultrahigh resolution mass spectrometry have provided new opportunities to study the interactions *in situ*.

The diversity of pelagic bacteria is strongly correlated to latitude and temperature, weaker correlations were detected to productivity-related parameters based on the analysis of 103 globally distributed community fingerprints (Fuhrman et al. 2008). At a much higher resolution based on more than 500 samples throughout various environments, distinct communities were detected for surface, deep water, coastal, open ocean and benthic habitats, confirming latitudinal trends and unraveling vertical trends (Zinger et al. 2011). But even on the smaller scales of an estuary, a succession of the microbial community succession could be followed along the salinity gradient (Crump et al. 2004; Fortunato et al. 2012). From Arnosti et al. (2011) we learn that a latitudinal gradient not only exists for the microbial communities, but is also reflected in the range of complex substrates that are available and degradable by the respective community. Most studies, however, focus on trends in DOM composition over the seasonal variation or the change in DOM lability with depth. The annual new production in the photic zone leads to an increased DOC concentration in the upper water layers (Carlson et al. 1998; Jiao and Zheng 2011; Ogawa and Tanoue 2003) and the more recently produced DOM can be distinguished from aged DOM by established methods such as the amino acid composition (Dauwe and Middelburg 1998) or radiocarbon dating (Flerus et al. 2012).

Although DOM and microbial community composition have no essential need to correlate due to the high genetic potential of the bacterial populations (Wilmes and Bond 2009) and the capability to perform lateral gene transfer (Cottrell et al. 2000), studies nevertheless suggest an important coherence. The *amount* of bioavailable DOM fuels the microbial diversity (Landa et al. 2013), but also the *composition* of the DOM plays a role. According to Pinhassi et al. (1999), a shift in microbial community composition occurred in incubation experiments after the addition of protein. Similarly, Kirchman et al. (2004) report that DOM from different sources

can affect the abundance of major bacterial groups (*Alphaproteo*-, *Betaproteo*-, *Gammaproteo*-, *Cytophaga*-like bacteria), which in turn have different capabilities to impact DOM hydrolysis. This phenomenon has also been observed *in situ* – *Bacteroidetes*, *Gammaproteobacteria* and *Alphaproteobacteria* consecutively constituted the most abundant bacterial taxa during the degradation of organic matter after an algal bloom (Teeling et al. 2012). *Alphaproteobacteria*, especially the ubiquitous SAR11 clade, seem to be well adapted to the uptake of amino acids and sugars (Malmstrom et al. 2004), while e.g. Archaea assimilate L-aspartic acid faster than Bacteria (Teira et al. 2006).

Over the last decades, the understanding of microbe-DOM interactions have greatly improved and continue to deepen, but we still need to learn more about the detailed relationships on molecular level, and on both sides – the microbial community and the DOM composition. One of the most prominent examples of our lack of knowledge and need of further research is the fact that huge amounts of carbon reside in the DOM pool for thousands of years in the deep ocean and are not metabolized by the microorganisms, in spite of containing a perfectly suitable substrate including essential elements like C, H, O, N, P, and S. A disruption of the balance of carbon bound in CO_2 and fixed in organic matter through an increase in DOM uptake by 1% would yield a CO_2 flux into the atmosphere equal to the annual global CO_2 emission by fossil fuel combustion – a process with unknown impact on the global carbon cycle (Nagata 2008).

1.1.4 Current issues in DOM research

Researchers have over the last century recognized the high significance of DOM for the global carbon cycle and gained insight into most major sources and sinks. The detailed mechanisms behind DOM cycling and pathways through the oceans, however, remain to be unraveled.

Stability of DOM. Most of the organically bound carbon dissolved in the oceans resides in the DOM pool for thousands of years and is not metabolized by the

microbes (Hansell 2013), despite the limited nutrient supply in the deep ocean. A role of microorganisms in the formation of this refractory material is assumed, but do they also play a role in its removal? Dittmar (2014) summarized mechanisms that might be, also in combination, responsible for the long-term stability of the oceanic DOM. The lack of essential minerals needed during the degradation may hinder the uptake (Kritzberg et al. 2010), or the molecules may possess an intrinsic stability protecting them from metabolization (Jiao et al. 2010). Furthermore is it possible that the diversity of the single compounds is too high (Dittmar and Paeng 2009; Hertkorn et al. 2007) and their respective concentrations too low to induce metabolic pathways. Abiotic processes such as photochemically induced transformations, adsorption to particles and burial in sediments have not yet been excluded as possible sinks of refractory DOM. Because of the long timescales and low changes in concentrations, these processes can hardly be studied under controlled laboratory conditions using marine DOM. Evidence of refractory, high molecular weight DOM turnover involving the addition of a labile substrate have however been obtained from lake DOM incubations (Geller 1985).

Quantification and structure determination. The advent of ultrahigh resolution techniques has brought the characterization of the complex mixture that is natural DOM a huge step forward. One disadvantage of the ultrahigh resolutions mass spectrometry via ESI FT ICR MS is its semi-quantitative nature. The use of internal standards is difficult due to the complex matrix as well as unknown influences on the ionization and thus so far not established for this method. However, the quantification of metabolites through the addition of isotopically-labeled internal standards has been successfully applied (Han et al. 2008). Moreover, FT ICR MS does not provide information on the structure of a compound, which may be crucial for the possible microbial uptake of the molecule (Onumpai et al. 2011). One molecular formula may represent thousands of structures with a characteristic distribution over environments (Hertkorn et al. 2013; Hertkorn et al. 2007), a diversity that remains hidden so far.

Fate of terrigenous DOM. Every year, 0.25 Pg C reach the world's oceans via rivers (Cauwet 2002). This DOM, despite its high molecular compliance with marine DOM (Koch et al. 2005), is generally regarded as more degraded and carries lignin-derived phenols as a tracer of vascular plant-derived material (Opsahl and Benner 1997). However, only traces of lignin phenols can be detected in most main water masses and the calculated residence times are rather short compared to bulk DOM. There is some evidence that riverine DOM provides valuable nutrients to otherwise potentially nutrient limited microbial communities in coastal zones (Tranvik 1992), but this factor is not sufficient to explain its ephemerality. Terrigenous DOM is usually more aromatic and thus more susceptible to photodegradation – which again does not provide an explanation for the terrigenous DOM turnover in the Arctic Ocean where the solar irradiation is low (Benner et al. 2005). The overall distribution and reactivity of terrigenous DOM, especially the potential microbially mediated turnover, need to be studied in more detail to understand its fate.

Environmental "omics" and synthesis of datasets. High throughput sequencing methods such as 454 pyrosequencing (Elahi and Ronaghi 2004) enable the large-scale characterization of microbial communities, which in combination with meta-transcriptome and -proteome studies allow the inference of the function of the community (Grossart 2011). Through these techniques, the microbial community can be analyzed at a resolution similar to or even superior to the molecular analysis of DOM via FT ICR MS. The generation of data is now possible, but the bioinformatics tools to intelligibly evaluate (Bakker et al. 2012) and combine the huge datasets in order to obtain the best possible knowledge are still in their infancy.

1.2 Objectives and outline of this thesis

The aim of this PhD work is the elucidation of the origin and fate of DOM in the ocean on molecular level. Specifically, the present study addresses the following questions regarding marine DOM:

- What is the diversity and composition of marine DOM on molecular level?
- What are the sources and sinks of refractory DOM?
- Can we trace terrigenous DOM input to the ocean?
- Which role do microorganisms play in the production and reshaping of the DOM pool?
- Can we identify common trends between microbial community and DOM molecular composition and identify key players?

Chapter 2 focuses on the fast turnover of DOM produced during the spring and summer phytoplankton blooms during the short productive season in the high Arctic. Chapter 3 describes a long-term laboratory incubation study, which followed the production and subsequent degradation of DOM on molecular level. Chapter 4 examines possible links between DOM and microbial community composition within the salinity gradient of the North Sea and introduces statistical methods to handle complex data sets.

1.3 Contributions to publications

This thesis includes the complete version of one published manuscript (chapter 2). Chapter 3 and Chapter 4 each include manuscripts in a form ready for submission. The content of the published manuscript is unchanged but the style adapted to the format of this thesis.

Chapter 2 - Molecular evidence for rapid dissolved organic matter turnover in Arctic fjords (published manuscript)

Helena Osterholz, Thorsten Dittmar, Jutta Niggemann

The study was initiated and designed by J.N. and T.D.. H.O. and J.N. carried out field work and sample preparation. H.O. performed all microbiological analyses and ultrahigh resolution mass spectrometry, J.N. did dBC analysis. All data analysis including statistical analysis was done by H.O., advised by J.N. and T.D.. All authors contributed to data interpretation and general discussion. H.O. wrote the manuscript with input from J.N. and T.D..

published in *Marine Chemistry*, vol 160, pages 1-10, doi: 10.1016/j.marchem. 2014.01.002

Chapter 3 – Do marine microorganisms really produce refractory dissolved organic matter? (manuscript in preparation)

Helena Osterholz, Jutta Niggemann, Helge-Ansgar Giebel, Meinhard Simon, Thorsten Dittmar

TD, JN, HO and MS conceived the study, HO performed all analyses and data handling. HAG performed flow cytometric analysis and enumeration of microorganisms. All authors contributed to data interpretation. HO wrote the manuscript with significant input from all coauthors.

This manuscript is intended for submission to Nature and therefore prepared in the journal's style.

Chapter 4 - Terrigenous input and microbial processing - driving forces of dissolved organic matter composition in the North Sea (manuscript in preparation)

Helena Osterholz, Gabriel Singer, Jutta Niggemann, Meinhard Simon Thorsten Dittmar

This manuscript includes data acquired during RV Heincke cruise 361 by the participants from the ICBM. H.O. took the samples for DOM characterization and performed FT-ICR-MS analyses. G.S. provided expertise in multivariate analyses combining microbial community and DOM datasets. J.N. and T.D. aided with data evaluation and interpretation. H.O. wrote the manuscript with editorial input from all co-authors.

The manuscript is intended for submission to The ISME Journal.

2 Molecular evidence for rapid dissolved organic matter turnover in Arctic fjords

Helena Osterholz, Thorsten Dittmar, Jutta Niggemann

published in *Marine Chemistry*.

vol. 160, pages 1-10, doi: 10.1016/j.marchem.2014.01.002

Max Planck Research Group for Marine Geochemistry, Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky University Oldenburg, Carl-von-Ossietzky-Strasse 9-11, 26129 Oldenburg, Germany

2.1 Highlights

- DOM produced during spring and summer plankton blooms is efficiently recycled.
- Autochthonous production does not leave persistent imprint in Arctic DOM pool.
- In late summer, DOM concentration in Svalbard fjords was uniformly low.
- Molecular DOM composition was indistinguishable among locations and water depths.
- Microbial community composition changed with location and water depth.

2.2 Abstract

Dissolved organic matter (DOM) in the ocean comprises one of the largest active carbon pools on earth. Deep water formation at high latitudes carries DOM from the active surface layers to the deep ocean. However, information on sources and fate of DOM in the Arctic Ocean is limited. To reveal the relevance of autochthonous DOM production and transformation in Arctic fjord systems to the global deep ocean DOM pool, we performed a comprehensive study on the molecular composition of DOM and the composition of the associated microbial communities in four selected fjords of Svalbard. At various water depths, a total of 34 samples were taken in fall 2010 for the determination of bulk concentrations of dissolved organic carbon (DOC) and total dissolved nitrogen (TDN), for the molecular characterization of solid-phase extractable DOM as well as for microbial community fingerprinting.

While TDN concentration and the composition of the microbial community showed a clear distinction between surface and bottom water samples, bulk DOC (~60 µmol C L^{-1}) and dissolved black carbon (~1.8% of DOC) as a marker for terrestrial input were uniformly distributed. In-depth molecular-level analyses of the DOM composition using ultrahigh resolution mass spectrometry via Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) revealed insignificant variation of the relative abundance of 11630 molecular masses that were detected in the water samples.

From these findings we conclude that DOM produced during the spring/summer bloom is rapidly transformed within the short, but productive warm season by the specialized resident microbial community. Thus, in fall the DOM pool mainly consists of semi-refractory and refractory material, most of which has been introduced from Arctic Ocean water inflow. Assuming that our findings are representative for high latitude marine systems in general, the contribution of autochthonous seasonal DOC production in plankton bloom situations to the DOC pool in regions of deep water formation might be marginal.

2.3 Introduction

Marine dissolved organic matter (DOM) comprises one of the largest active carbon pools on earth, similar in size to atmospheric CO_2 or all land plant biomass (Hedges 1992). The global oceanic net primary production is estimated at 50 Gt C per year (Hedges 1992; Williams and Druffel 1987). A significant fraction of the fixed carbon is transferred to the DOM pool, where it is readily remineralized by microorganisms within the microbial loop (Azam 1998; Jiao et al. 2010; Pomeroy 1974) or channelled into the deep ocean, where it resides for hundreds to thousands of years.

Hansell (2013) divided the oceanic DOM pool into five fractions with distinct lifetimes along a continuum of reactivity starting with labile DOM that is removed within hours to days and thus does not accumulate in seawater. The observable, recalcitrant DOC fractions are further distinguished as semi-labile DOM with turnover times between months and years, which can be followed as seasonal DOC variation in the surface ocean, semi-refractory DOM with a model lifetime of about 20 years, refractory DOM with a model lifetime of 16,000 years, and the ultra-refractory DOM with a lifetime of around 40,000 years. Most DOC is retained in the refractory DOM pool in the deep ocean, comprising ~630 Gt carbon. The reasons why so much DOM escapes the microbial respiration even though it constitutes the major carbon and energy resource to support bacterial life are not yet well understood (Dittmar and Paeng 2009; Jiao et al. 2011).

For a long time, focus in DOM research has been on the analysis of selected compound classes in DOM, e.g. amino acids, carbohydrates (Amon et al. 2001), lignin (Opsahl et al. 1999) or dissolved black carbon (DBC, Dittmar 2008). Together these groups account for only ~5% of the total DOM, however, thus ignoring the major share. Only recent advances in technology, especially of ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS), have made it possible to molecularly characterize the compounds making up this huge DOM reservoir on a molecular level. Combined with soft electrospray ionization (ESI), intact polar molecules can be analysed, and the high mass accuracy allows for the assignment of molecular formulae to >70% of the tens of thousands of detected peaks per sample (Stenson et al. 2003). Previous research on the molecular composition of DOM by FT-ICR-MS revealed a universal fingerprint consisting of apparently biologically refractory compounds that is found all over the world's oceans (Dittmar and Paeng 2009; Flerus et al. 2012). Other examples of successful FT-ICR-MS applications include the differentiation of terrigenous and marine DOM (Koch et al. 2005), the characterization of biodegradable DOM in rivers (Kim et al. 2006), the identification of possible source markers for photodegradation and bacterial alteration (Kujawinski et al. 2009), and the tracing of the age of bulk marine DOM by combining FT-ICR-MS measurements with radiocarbon age determination (Flerus et al. 2012).

The Arctic is, besides the Southern Ocean, a major site of ocean deep water formation (Rudels and Quadfasel 1991). Thus, Arctic waters are a major source of DOM to the deep ocean (Amon et al. 2003), but reports on DOM concentration and especially composition in the Arctic Ocean are scarce. Previous studies focused on input and distribution of terrigenous DOM by the major Russian and Canadian rivers (e.g. McClelland et al. 2012). Here, in order to reveal the relevance of autochthonous DOM production and transformation in Arctic fjord systems, we performed a comprehensive study on the molecular composition of DOM and the composition of the simultaneously sampled microbial community in four fjords of Svalbard. The archipelago of Svalbard is located in the high Arctic between 76 and 81° north latitude and 10 to 30° east longitude (Fig. 1). The fjords on the western

coast are influenced by Atlantic and Arctic water masses and the glacial meltwater inflow at their heads (Hop et al. 2002). Among them, the Kongsfjorden ecosystem is the best-studied area with respect to phytoplankton dynamics, microbial communities and oceanography (Cottier et al. 2005; Hop et al. 2002; Iversen and Seuthe 2011; Svendsen et al. 2002).

Anthropogenic influence on the remote ecosystem of Svalbard is marginal, and the fjord water masses are well stratified (Svendsen et al. 2002). Thus, Svalbard fjords are ideal model systems to study the differences between fresher surface water that has recently experienced primary production and the older bottom waters. We hypothesized that glacial meltwater input and phytoplankton blooms affect surface water DOM composition and the resident microbial community, and that this imprint is clearly distinguishable when compared to bottom waters. In order to characterize the molecular composition of the fjord DOM, we applied ultrahigh resolution mass spectrometry via FT-ICR-MS. In addition, dissolved black carbon (DBC) was quantified as a molecular marker to assess land-derived DOM, which has been shown to carry a universal thermogenic imprint worldwide (Jaffé et al. 2013). Surface and bottom water microbial communities were characterized by total bacterial cell counts using flow cytometry and fingerprinting via denaturing gradient gel electrophoresis (DGGE).

2.4 Materials & Methods

Site description and sampling. During the field campaign (August 28 to September 5, 2010), the water masses of four fjords on the west coast of Svalbard, Norway, were sampled (Table 1, Fig. 1). Vertical profiles of temperature and salinity were recorded by a CTD (Seabird MicroCAT-SBE-37) for three stations per fjord located on transects from fjord head towards the ocean. Three to four water depths were chosen at each station based on the CTD profiles: surface water was taken with a bucket, intermediate water samples from the pycnocline and deep water samples close to the bottom were retrieved with a 4 L Niskin bottle. Subsamples for total cell counts were preserved with glutaraldehyde (25%, Carl Roth, Germany) at a final

concentration of 1% and stored at -20°C. Subsamples for molecular microbiological analysis were filtered on board and stored at -20°C. The water samples for chemical DOM analysis were stored at ambient temperature (~0°C) in the dark and processed on board or upon arrival at Kings Bay Marine Laboratory, Ny Ålesund, within 48 hours.

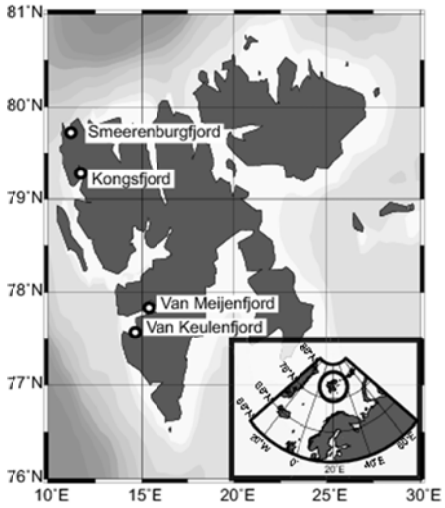

Fig. 1. Map of Svalbard, showing locations of sampled fjords. The rectangular inlay shows the location of the Svalbard archipelago (circled) north of the European mainland. Map created using Ocean Data View (R. Schlitzer, http://odv.awi.de).

A small meltwater stream and a piece of floating iceberg (cleaned by repeated rinsing with ultrapure water and melted in the laboratory) were sampled in the vicinity of Ny Ålesund to assess the input of DOC, TDN and DBC by land-derived freshwater input.

Bacterial cell counts (flow cytometry). The fixed samples were thawed and stained with SybrGreen I (1:10,000 dilution of commercial stock, Molecular Probes, United Kingdom) for 30 min at room temperature. After the addition of a defined amount of a polychromatic bead suspension (0.5 µm Fluoresbrite Microspheres, Polysciences Inc., USA), the determination of cell numbers was conducted with a FACScalibur flow cytometer (Becton Dickinson, USA). The flow rate was determined using the suspension of polychromatic beads before and after analysis (Marie et al. 1997). Data acquisition and analysis were performed with the Cell Quest Pro software (Becton Dickinson, USA).

Table 1: Sampling sites by fjord, station name, coordinates, water depth, depth group, temperature and salinity of sampled waters.

Fjord	Station	Coordinates	Depth [m]	Depth Group	Temperature [°C]	Salinity
Van Keulenfjord	HA	77°32.02 N 15°50.31 E	0	surface	1.5	31.6
			4	intermediate	1.5	32.4
			20	deep	1.2	32.8
	AC	77°33.29 N 15°39.29 E	0	surface	1.8	27.0
			5	intermediate	1.9	27.4
			40	deep	0.7	33.3
	AB	77°35.19 N 15°05.24 E	0	surface	3.2	32.6
			5	intermediate	2.6	33.0
			86	deep	0.5	33.8
Van Meijenfjord	AH	77°45.75 N 15°03.23 E	0	surface	4.3	29.9
			5	intermediate	3.5	31.5
			60	intermediate	1.0	32.8
			108	deep	-1.2	33.5
Kongsfjord	Q	78°59.43 N 12°17.87 E	0	surface	1.2	32.0
			10	intermediate	1.1	32.1
			40	deep	0.8	33.3
	F	78.54.97 N 12°16.04 E	0	surface	2.2	30.3
			8.5	intermediate	2.2	30.4
			69	deep	0.1	33.8
	P	78 54.47 N 12 10.16 E	0	surface	1.3	30.9
			5	intermediate	1.3	31.8
			57	deep	-0.4	33.8
	HB	79°02.09 N 11°42.12 E	0	surface	3.5	31.0
			4.5	intermediate	3.9	32.1
			195	deep	1.2	34.0
Smeerenburg-fjord	GK	79°38.49 N 11°20.97 E	0	surface	2.7	32.9
			12	intermediate	2.6	33.3
			165	deep	1.8	33.8
	J	79°42.74 N 11°05.19 E	0	surface	2.9	33.7
			8	intermediate	2.9	33.2
			200	deep	1.0	33.7
	GN	79°45.08 N 11°05.18 E	0	surface	3.0	31.0
			20	intermediate	2.8	33.4
			181	deep	0.7	33.7

Microbial community analysis (DGGE). Samples taken in Smeerenburgfjord and in van Keulenfjord were prepared for microbial community analysis. 500 mL of seawater were filtered through 3 µm polycarbonate (PC) filters (Millipore, USA) to retain the particle-associated bacteria. Of this filtrate, 250 mL were passed through 0.2 µm PC filters (Millipore, USA) to retain the free-living microbial community. The filters were stored frozen at -20°C. The DNA extraction procedure was modified after Zhou et al. (1996). The bacterial primer set 907R (5´- CCG TCA

ATT CM TTT GAG TTT - 3′) and GM5F (5′- (CGC CCG CCG CGC CCC GCG CCC GTC CCG CCG CCC CCG CCC G)-CC TAC GGG AGG CAG CAG - 3′, both MWG Biotech, Germany) was used for amplification of 16S rRNA gene fragments in a touchdown PCR program modified after Muyzer et al. (1993). DGGE was performed on an Ingeny phor U 2x2 (Ingeny International BV, Netherlands) system with a gradient from 40 to 70% urea/formamide at a constant voltage of 100 V for 20 h in TAE buffer (40 mmol L^{-1} Tris-acetate, 1 mmol L^{-1} EDTA, pH 7.4). Cluster analysis of DGGE band pattern using densitometric curves was carried out with the Gel ComparII software package (Similarity: Pearson Correlation, Clustering Method: UPGMA, Version 6.5, Applied Maths, Belgium).

Dissolved organic matter. Water samples for DOM extraction were filtered through glass fiber filters, using a 2 µm filter on top of a 0.7 µm filter (GMF and GF/F, Whatman, United Kingdom, combusted 400°C, 4 h) and acidified to pH 2 (HCl 25% p.a., Carl Roth, Germany). An aliquot of this acidified 0.7 µm filtrate was sampled for quantification of dissolved organic carbon (DOC) and total dissolved nitrogen (TDN). Additional duplicate samples for DOC and TDN were prepared from unfiltered water using GHP syringe top filters (0.2 µm, GHP Acrodisc, PALL Life Science, USA). Determined DOC and TDN concentrations of GF/F and GHP filtrates were without systematic trends (average difference is 3.5%, thus within analytical precision, see below). Reported concentrations are averaged values from triplicate samples.

DOC and TDN concentrations were analyzed by high temperature catalytic combustion using a Shimadzu TOC-VCPH/CPN Total Organic Carbon Analyzer equipped with an ASI-V autosampler and a TNM-1 module. Prior to analysis, the acidified samples were purged with synthetic air to remove dissolved inorganic carbon. L-arginine solutions ranging from 5 to 500 µmol C L^{-1} and 6.6 to 333.3 µmol N L^{-1}, respectively, were used for calibration and Deep Atlantic Seawater reference material (DSR, D.A. Hansell, University of Miami, Florida, USA) was measured during each run to control for instrumental precision and accuracy. Except for two obvious outliers, the standard deviation of triplicates was

always below 2.9 µmol C L^{-1} for DOC and 1.2 µmol N L^{-1} for TDN analysis. Precision was 3.6±1.8% for DOC and 2.2±0.4% for TDN, while accuracy was 4.3±5.3% for DOC and -6.6±2.8% for TDN, respectively.

DOM was extracted from 2 liters of filtered and acidified water with commercially available modified styrene divinyl benzene polymer cartridges (PPL, Agilent, USA) as described in Dittmar et al. (2008). After extraction, cartridges were rinsed with acidified ultrapure water (pH 2, HCl 25%, p.a., Carl Roth, Germany) to remove remaining salts, dried by flushing with nitrogen gas and eluted with 6 ml of methanol (HPLC-grade, Sigma-Aldrich, USA). Extracts were stored in amber vials at -20°C. The extraction efficiency was 61.3±4.7% on a carbon basis for fjord water DOM, similar to published extraction efficiencies for seawater DOM extraction on PPL resins (Dittmar et al. 2008). In previous studies, the same extraction method and sample preparation has successfully been applied to trace the aging of oceanic DOM by molecular-level characterization via ultrahigh resolution FT-ICR-MS (e.g. Dittmar and Paeng 2009; Flerus et al. 2012). Thus, the extracted DOM fraction should provide a representative subsample of the prevailing DOM independent of source and degradation stage.

Dissolved black carbon (DBC) was quantified in aliquots of the methanol extracts according to Dittmar (2008). In short, an equivalent of 2 µmol C per extract was oxidized with nitric acid at 170°C for 9 hours in sealed glass ampules placed in high pressure reaction vessels. After the oxidation, the samples were dried by evaporation of the nitric acid and redissolved in phosphate buffer (Na_2HPO_4 and NaH_2PO_4 each 5 mM in ultrapure water, pH 7.2). Separation of the benzenepolycarboxylic acids (BPCAs) was achieved on an ultra performance liquid chromatography system (Waters Acquity UPLC) using a BEH C18 Column (2.1×150 mm, 1.7 µm, Waters) with an aqueous phase/methanol gradient modified after Dittmar (2008). The aqueous phase consisted of a tetrabutylammonium bromide solution (4 mM) in phosphate buffer (Na_2HPO_4 and NaH_2PO_4 each 5mM in ultrapure water, pH 7.2). The BPCAs were identified according to retention time and absorbance spectra (220 to 380 nm). Quantification was performed using the adsorption signal at 240 nm and

external calibration with 4-point-standard curves covering the respective concentration range. The concentration of DBC was calculated from the concentrations of the individual BPCAs according to Dittmar (2008). Standard deviation of triplicate analysis was <5% for concentrations of individual BPCAs and DBC, and detection limit for DBC was 0.2 μmol L^{-1} seawater.

The mass spectra were obtained on a 15 Tesla Solarix FT-ICR-MS (Bruker Daltonics, USA) equipped with an electrospray ionization source (Bruker Apollo II) applied in negative mode. DOM extracts were diluted to a final DOC concentration of 10 mg C L^{-1} in a 1:1 mixture of ultrapure water and methanol (HPLC-grade, Sigma-Aldrich, USA). A total of 500 scans were accumulated per run, the mass window was set to 150-2000 Da. The spectra were calibrated with an internal calibration list using the Bruker Daltonics Data Analysis software package. The mass to charge, resolution and intensity were then exported and processed using in-house Matlab routines. Molecular formulae were assigned to peaks with a minimum signal-to-noise ratio of 4 following the rules published in Koch et al. (2007). For each molecular fomula, the double bond equivalents (DBE, DBE=1+½(2C-H+N+P)) as a measure for the degree of unsaturation and the Aromaticity Index (AI=(1+C-O-S-½H)/C-O-S-N-P)) as well as the modified Aromaticity Index (AI_{mod}=(1+C-½O-S-½H)/(C-½O-S-N-P)) to assess the presence and extent of aromatic structures (Koch and Dittmar 2006) were calculated.

North Pacific Deep Water reference. The North Pacific Deep Water (NPDW) is one of the oldest water masses on earth (Stuiver et al. 1983) and considered to be dominated by refractory DOM (e.g. Hansell 2013). To produce the in-house reference, large volumes of water from the North Pacific oxygen minimum zone at 670 m depth near Big Island (Hawaii) were extracted using the same extraction method described above for the Svalbard fjord samples.

Statistical analysis of FT-ICR-MS data. The 34 fjord samples were analyzed on the FT-ICR-MS alternating with the NPDW DOM reference and a mixed sample reference consisting of equal parts of the 34 fjord extracts. This analysis scheme allowed to assess and account for FT-ICR-MS instrument variability during the time

span needed for analysis of the complete sample set (4 days), thereby providing a robust basis for in-depth statistical analysis. Spectra were normalized to the sum of peak intensities per sample. For a more conservative approach, and in addition to the threshold of S/N 4, molecular masses were only included when they occurred at least three times in one of the groups (NPDW reference, mixed fjord sample reference, fjord samples). This removed rare and uncertain peaks. Three fjord samples were excluded from further analysis because of obvious contamination of the DOM extracts (station Q at 0 m, station HA at 4 m, station AC at 5 m).

FT-ICR-MS data was subjected to principal component analysis (PCA) conducted with the software The Unscrambler X 10.2 (validation method: cross validation, sample weights: 1, CAMO Software AS, Norway). For a first PCA aiming at the separation of Svalbard fjord DOM and NPDW DOM, only assigned molecular formulae present in all Svalbard fjord samples and the NPDW DOM reference were included. A second PCA intended to visualize the differences among the Svalbard fjord samples was based on all detected masses of the fjord samples.

We further compared the multivariate variability (dispersion) of the fjord samples with the variability due to instrumental error, using a dissimilarity-based approach. From a matrix of Euclidean distances computed from scaled (normalized and standardized) data, we calculated for each group the average distance to each group centroid. The average distance to the centroid is equivalent to the average distance among all pairwise group member combinations and serves as a measure of dispersion. Differences in dispersion between groups were tested for significance by permutation (Anderson 2006).

2.5 Results

Irrespective of location and water depth, DOC and DBC concentrations were similar in all 34 fjord water samples and averaged 59.5±1.7 µmol DOC L^{-1} and 1.8±0.1% DBC of DOC, respectively (Fig. 2). TDN concentrations were significantly higher in the deep fjord samples (surface 7.8±1.0 µmol L^{-1}, deep 12.6±2.2 µmol L^{-1}), while total bacterial cell counts (surface $2.25x10^6$±$0.78x10^6$ cells mL^{-1}, deep

$1.23x10^6 \pm 0.28x10^6$) were significantly lower in the deep fjord samples than at the surface (Mann-Whitney Rank Sum Test, $p<0.01$, Fig. 2).

The cluster analysis of the DGGE band pattern revealed differences in the microbial community composition of the stations (Fig. 3). The two main clusters separate free-living (0.2-3 µm) and particle-associated (>3 µm) bacteria. Minor clusters divide surface and bottom water samples, and the microbial community is more closely related within the same fjord. The clustering was more expressed for the free-living than for the particle-associated microbes.

A total of 11630 resolved molecular masses of singly charged compounds were detected in the FT-ICR-MS spectra, covering a mass range of 159-920 Da. A Constrained Analysis of Principal Coordinates (CAP, Oksanen et al. 2012) was conducted in order to reveal correlations between the environmental parameters and the relative intensities of the measured masses. Neither DOC, TDN, DBC and bacterial cell counts, nor sampling related parameters like extraction efficiency or time lag between sampling and sample processing significantly correlated with the molecular DOM composition.

From a first visual inspection, the mass spectra of the fjord samples exhibited a great similarity to those of the NPDW DOM reference material analyzed repeatedly (Fig. 4), even when comparing the spectra obtained for a single nominal mass (Fig. 4, top right). However, thorough visual inspection of the mass spectra revealed minor changes in presence and abundance of individual peaks (Fig. 4, bottom right).

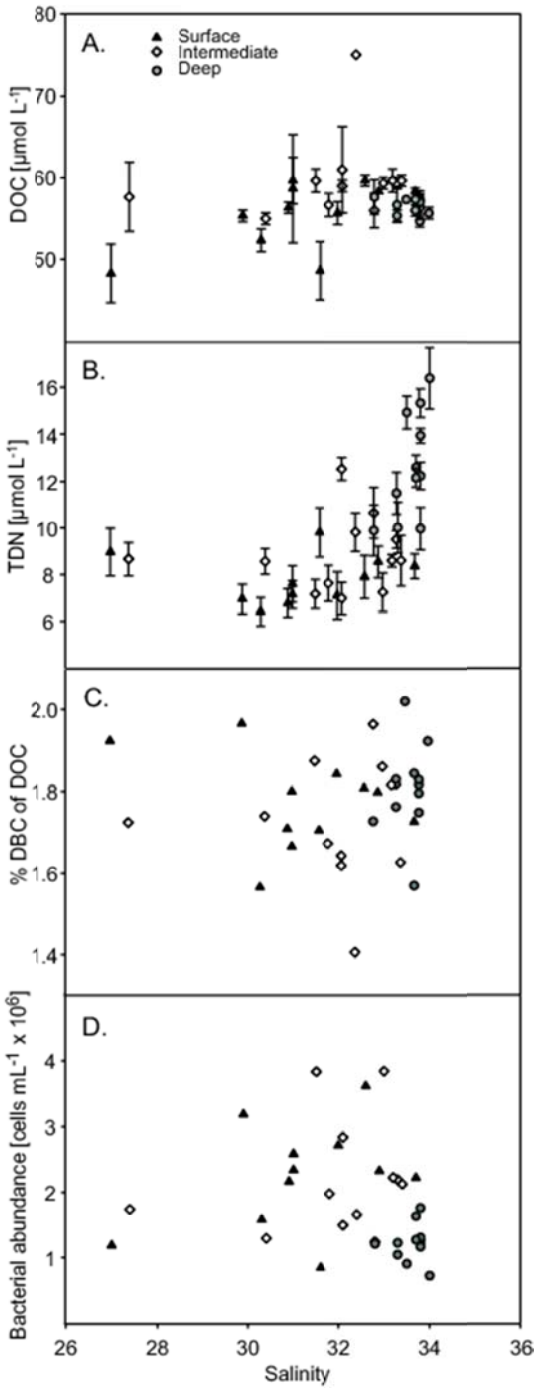

Fig. 2. Concentrations of A. dissolved organic carbon (DOC, average±s.d.; s.d. = standard deviation of triplicates), B. total dissolved nitrogen (TDN, average ± s.d.), C. % dissolved black carbon (DBC) of DOC and D. bacterial cell counts of all sampled Svalbard fjordwaters against salinity and separated by surface, intermediate and deep sampling depth. Error bars indicate standard deviation of triplicate measurements where applicable.

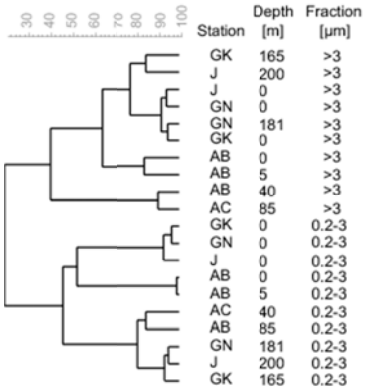

Fig. 3. Cluster dendrogram based on densitometric curves from denaturing gradient gel electrophoreses (DGGE) banding patterns for the free-living (0.2–3 µm fraction) and particle-associated (>3 µm fraction) microbial communities at all water depths in Smeerenburgfjord (GK, J, GN) and Van Keulenfjord (AB, AC).

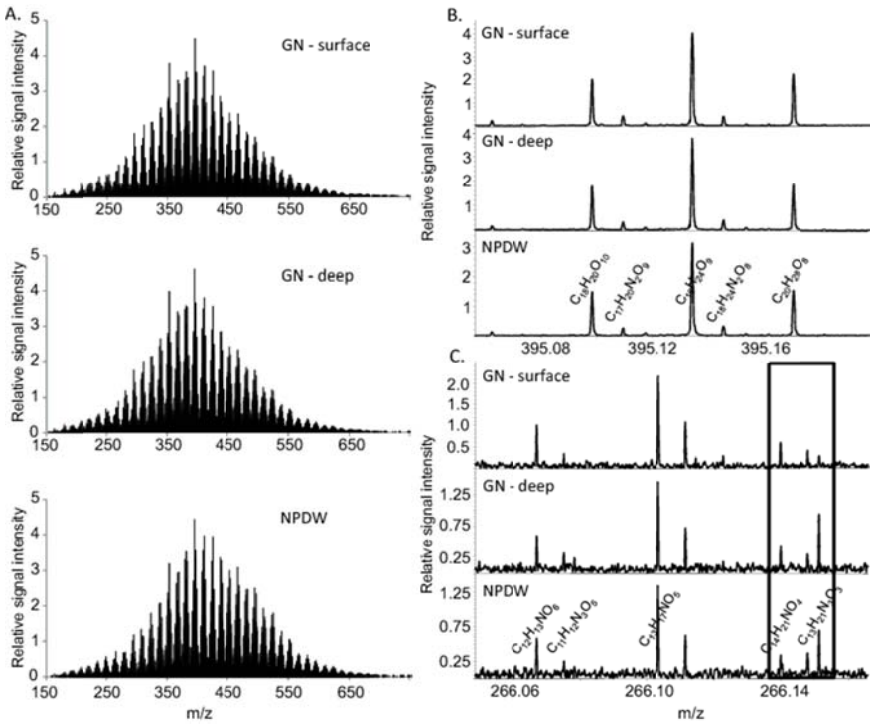

Fig. 4. A. Examples of Fourier-transform ion cyclotron resonance mass spectra of Svalbard surface water (Station GN, 0 m), Svalbard deep water (Station GN, 181 m) and the NPDWreference material, mass range from150 to 750 Da on the left and zoomed into the nominal masses B. 395 Da and C. 266 Da. Highlighted area marksmasses with changes in relative intensities.

Principal Component Analysis (PCA) of a reduced data set including only those masses that were present in all NPDW reference measurements and in all fjord samples (n=1782) separates both groups on PC1 which accounts for 84% of the variability in the data set (Fig. 5A). A PCA including only the fjord samples, and all detected masses of the fjord samples (n=11630, Fig. 5B) shows little explained variability for PC1 and PC2 with 10 and 7%, respectively, and does not reveal any separation of the samples by water depth. None of the principal components points to dominant influence of location (the different fjords), DOC concentration or any of the determined parameters (salinity, temperature, depth, TDN or DBC concentration, microbial cell counts).

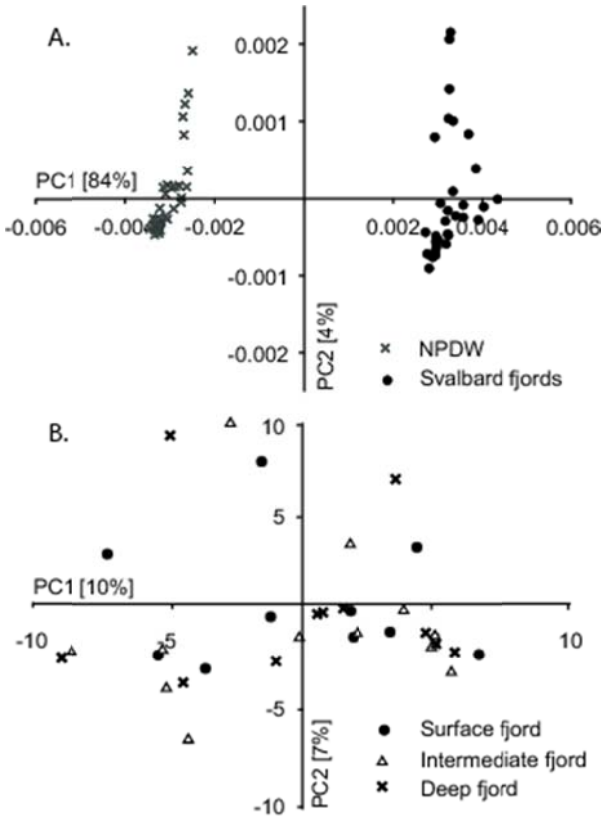

Fig. 5. A. Principal Component Analysis (PCA) of molecular masses shared between Svalbard fjord samples and NPDWDOM with assigned molecular formulae (n = 1782). B. Principal Component Analysis of all molecular masses detected in the fjord samples (n= 11630).

Differences between Svalbard fjord DOM and NPDW DOM are reflected in the comprehensive description of the molecular composition provided by the indices informing about molecular structure (Tab. 2). In general, the Svalbard samples are slightly more diverse with a higher number of detected masses, and the detected masses are on average at least 20 Da smaller than the masses detected in the NPDW DOM. The lower average double bond equivalent (DBE) of the fjord samples suggests a higher degree of saturation, while the aromaticity index (Dittmar and Koch 2006) does not differ between sample types. The average distance of each reference sample to the respective group centroid (Fig. 6, NPDW reference, fjord mixed reference SvalMix), the dispersion of the molecular data, reflects analytical variability. The dispersion of molecular data from a group of different samples (Sval, SvalSurface, SvalDeep) additionally represents differences in DOM composition of the individual samples. The dispersion of the repeated NPDW reference measurements is significantly lower than that of all fjord samples and the subgroup of deep fjord samples. However, the dispersion of the repeated fjord mixed sample measurements is not significantly lower than the fjord samples, indicating that the natural variability among samples was small and within the range of our analytical uncertainty (Fig. 6).

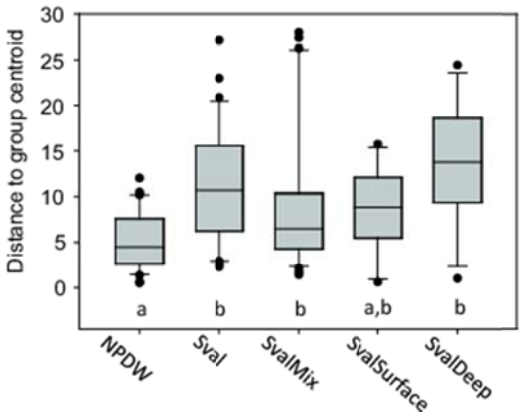

Fig. 6. Boxplots comparing the variation of DOM molecular composition (dispersion) across various groups. Each boxplot shows the distribution of the Euclidean distances of all samples to the centroid of the respective group. NPDW (n = 35), Sval - all fjord samples (n = 31), SvalMix — all samples mixed in equal parts (n = 28), SvalSurface - allfjord surface samples (n = 10), SvalDeep - all fjord bottom water samples (n = 11). Letters a and b denote significant differences between tested groups (p b 0.05).

Table 2: Comparison of FT-ICR-MS spectra, summarizing all Svalbard fjord samples (n = 31) and all replicate analyses of the NPDW (North Pacific Deep Water) DOM reference (n = 34). All numbers are averages (weighted by FT-ICR-MS peak intensity), s.d. = standard deviation, MF = molecular formulae.

	Svalbard (average±s.d.)	NPDW (average±s.d.)
General		
number of peaks	6087(±634)	5787(±272)
number of assigned MF	3653(±227)	3638(±107)
% of masses with assigned MF	60(±3)	63(±1)
average m/z of all peaks	410.4(±2.4)	434.8(±2.2)
Elemental Composition		
average C	19.63(±0.10)	20.66(±0.09)
average H	24.85(±0.14)	25.97(±0.11)
average O	8.92(±0.08)	9.41(±0.05)
average N	0.41(±0.01)	0.41(±0.01)
average P	0.04(±0.00)	0.05(±0.00)
average S	0.10(±0.01)	0.13(±0.00)
average O/C	0.46(±0.00)	0.46(±0.00)
average H/C	1.27(±0.00)	1.26(±0.00)
Molecular Indices		
average DBE	8.43(±0.05)	8.91(±0.04)
average AI	0.20(±0.00)	0.20(±0.00)
average AI_{mod}	0.26(±0.00)	0.26(±0.00)

2.6 Discussion

Characterization of prevailing water masses. During the sampling campaign, the water columns of the fjords were partially stratified according to gradients in salinity, temperature and TDN concentrations. The salinity gradient was most pronounced in the southernmost Van Keulenfjord with a surface water salinity of 27.0 increasing to 33.3 at 40 m depth (station AC), indicating dilution of prevailing seawater with up to ~26% of freshwater, while Kongsfjord and Smeerenburgfjord in the North showed almost no salinity gradient. In all fjords, the comparatively low water temperatures (-1.2 to 1.8°C) and constantly high salinities (33.3-34.0) of the deep water layers indicate influence of Transformed Atlantic Water (T>1°C, S>34.7) which originates from mixing of Arctic Water (>3°C, S>34.9) transported

by the West Spitsbergen Current with colder, less saline Arctic Water. The even lower salinities of the surface waters reflect freshwater input to the fjords (Svendsen et al. 2002). Cottier et al. (2005) report increased contribution of Atlantic Water in the Kongsfjord over summer months, with a cumulative annual freshwater input (glaciers, snowmelt, precipitation, runoff, groundwater discharge) accounting for up to 5% of the total fjord volume of 29.4 km^3. Thus, lowest surface water salinity and strongest water column stratification occur during peak freshwater input, i.e. times of snowmelt during early summer.

In our study, the chemical parameter most suitable to trace the recent history of the water masses in the fjord is the concentration of TDN. The surface nitrogen pool is partially replenished by the freshwater input (8.2-9.0 µmol L^{-1} TDN measured in the freshwater and melted iceberg samples), but TDN concentrations were generally higher in the deeper water layers where no primary production takes place. The inorganic nitrogen in the photic zone is an essential nutrient and used by the plankton especially during spring and summer blooms to build proteins, nucleic acids and other cell parts. Phytoplankton blooms, the major source of autochthonous labile DOM to the fjord system, occur from March to June, peaking around April (Eilertsen et al. 1989; Iversen and Seuthe 2011). Wängberg et al. (2008) report DOC concentrations ranging from 63 to 73 µmol C L^{-1} in June in Kongsfjord, an up to 35% higher DOC concentration compared to our study in August, when DOC concentrations averaged 59.5±1.7 µmol DOC L^{-1} and were rather constant along the narrow salinity gradient (Fig. 2). The two potential freshwater endmembers, sampled near Ny Ålesund (a glacial meltwater river and a melted iceberg) were low in DOC (7.0-12.5 µmol C L^{-1}) and DBC (1.7-0.3 % BC of DOC), and would thus dilute the surface water DOC and DBC concentrations rather than serving as sources. Furthermore, it has been shown that glacial runoff contains mainly labile reduced carbon of microbial origin (Hood and Berner 2009), which can be degraded quickly in the fjords. However, the freshwater input at the time of the sampling campaign was low due to the cold temperatures and complete freezing of the small rivers only a few days after sampling.

Microbial community composition and rapid processing of fresh DOM. The prokaryotic abundance was significantly higher in surface waters than in the deep water layers (Fig. 2) and differences were also evident with regard to the microbial community composition: variability related to habitat (free-living or particle-attached), location (fjord) and water depth were revealed via DGGE (Fig. 3). Since DOM composition is indistinguishable whereas the microbial community can be discerned between sites and depths, the initial hypothesis that DOM composition and bacterial community structure are directly related cannot be confirmed in this case. Other environmental parameters regulating the microbial community structure must exist. While DOM did not retain a characteristic molecular signature of its recent history, the microbial community composition might still carry a specific imprint of the recent bloom events. Furthermore, the preference of or metabolic advantage at a certain temperature, light condition or nutrient availability might also be responsible for the observed differences in microbial community composition.

It is well established that the community composition of particle-associated bacteria differs from that of free-living bacteria (Fandino et al. 2001; Grossart et al. 2005; Stevens et al. 2005). Teske et al. (2011) also observed differences in community composition between surface and bottom water in Smeerenburgfjord (station J). Furthermore, these authors were able to show that despite the cold temperatures, efficient carbon cycling exists at high latitudes. This may be due to enzymes with low temperature optima adapted to the Arctic conditions (Arnosti and Jørgensen 2003, Arnosti and Steen 2013). Iversen and Seuthe (2011) report efficient DOC cycling following the spring blooms in the Arctic, and average turnover times of 0.1 to 1.1 years for semi-labile DOM in the Arctic were calculated by Wheeler et al. (1996). There is thus much evidence that bacterial activity is not necessarily depressed in cold Arctic waters and that it is likely that freshly produced DOM during phytoplankton blooms is rapidly turned over by the resident microbial community.

In order to test the plausibility of our results, we conservatively estimated whether the DOM that had accumulated during the spring/summer bloom can be consumed

within the short Arctic summer period. We assumed that the dissolved inorganic nitrogen (DIN) in the surface water resulted from DOM mineralization, implying that DIN was limiting primary production and thus depleted during the phytoplankton bloom. Since only TDN was measured, we estimated the DIN by applying the average C/N ratio of 13.4±6.6 for surface ocean DOM summarized by Bronk (2002), yielding a minimum surface DIN concentration of 1.6 to 3.2 μmol L^{-1}. The concentration may however be overestimated when inorganic nitrogen from the deep water was mixed into the surface (increasing TDN) or the C/N ratio of DOM was lower than the presumed range. The estimated amount of DIN in surface waters could be produced by the remineralization of 17-53 μmol L^{-1} DOC if only DOM with a C/N ratio of 13.4±6.6 (Bronk 2002) was taken up as a substrate. Kirchman et al. (1991) estimated DOC turnover rates during a spring phytoplankton bloom in the North Atlantic ranging from 0.006 to 0.1 d^{-1}. Applying these rates to a calculated maximum DOC concentration of 69-105 μmol L^{-1} (lowest measured DOC concentration of 52 μmol L^{-1} plus estimated remineralized DOC of 17 to 53 μmol L^{-1}), we found that the difference to the observed minimum DOC concentration of 52 μmol L^{-1} can be degraded within 3 to 86 days. In support of our findings, Wheeler et al. (1997) estimated that during the productive season in the central Arctic Ocean, bacterial consumption exceeds the in situ primary production of DOC. Kirchman et al. (1991) report slight increases in DOC concentration of 10-50 μmol L^{-1} during a spring bloom in the open ocean, and Duursma (1963) observed that DOC concentrations decreased from a bloom situation in spring by a third until winter. These observations together with the reports on well-adapted and efficient Arctic microbial communities are consistent with the proposed scenario of episodic DOM production during spring and summer phytoplankton blooms in Svalbard fjords and an almost instantaneous degradation of the introduced fresh DOM signature.

Molecular-level characterization of fjord DOM. We hypothesize that the consistently low DOC concentrations observed during our field study in August 2010 represent a typical early fall situation in Arctic fjords. Fresh labile DOM added by primary production during spring and summer plankton blooms is readily turned over and does not persist in the fjord waters. In support of this, the recent primary

production left no detectable imprint on the molecular DOM composition. While samples are not available to directly trace the rapid DOM turnover, our dataset includes the seasonal endmembers, i.e. high TDN deep water representing the pre-bloom condition and low TDN surface water representing the past-bloom conditions. Thus, the spatial resolution of the sampling provides information on the timescale of the proposed processes.

Overall, contributions of nitrogen-containing, peptide molecular formulae typically produced during plankton blooms and marker compounds for fresh organic matter (Berman and Bronk 2003; Singer et al. 2012) were not elevated in surface water samples and the composition of the surface fjord DOM with a history of recent primary production was indistinguishable from that of the deep fjord water samples. In accordance with this, extraction efficiencies were similar for all samples and in the range reported for refractory marine DOM (Dittmar et al. 2008; Flerus et al. 2012). Taking into account the distribution of all compounds as assessed via FT-ICR-MS, we find very low variation in the DOM pool without significant correlation to any of the known environmental parameters (salinity, depth, DOC, TDN or microbial cell counts), leading to the assumption that most of the labile and semi-labile DOM that was produced during the spring/summer blooms in the fjords was already transformed into semi-refractory/refractory DOM at the time of sampling. This is remarkable considering the multitude of degradation and transformation processes involved in the microbial recycling of DOM. Thus, the autochthonous organic matter added to the fjords during spring and summer phytoplankton blooms does not persistently affect the fjord DOM pool, neither quantitatively regarding the carbon budget nor qualitatively regarding the molecular composition.

The fjord DOM samples are still significantly different from NPDW DOM despite their similarity in most chemical parameters (Fig. 5, Tab. 2). The lower average mass of the fjord DOM molecules (410±4 Da, weighted average) and the higher average mass of the NPDW reference (434±8 Da) are consistent with the findings of Flerus et al. (2012), who describe a shift of the average molecule size from recent to

aged DOM. However, their values are generally lower, ranging from 408 Da for fresher to 418 Da for aged DOM on a transect in the eastern Atlantic Ocean. The smaller average size of the fjord molecules could be explained by input of terrestrial DOM (Amon and Meon 2004). The islands of Svalbard have a low plant species richness dominated by lichens and mosses with few higher plants (Elvebakk 1994), but the Arctic Ocean in total receives considerable amounts of terrestrial DOM via the large Russian and Canadian rivers (Dittmar and Kattner 2003; Wheeler et al. 1997). An extensive input of DBC originating from glacier abrasion and erosion of coal seams by rivers and meltwater from the islands of Svalbard could not be confirmed, and thus the DBC detected in the fjords must originate from Arctic Water flowing into the fjords.

On the timescale covered by this study (a few months), including pre-bloom (deep) and post-bloom (surface) water, there were no significant modifications of the DOM. However, the differences between the DOM from the Arctic and the North Pacific can be interpreted as a selective decomposition of specific compounds over long terms of several hundred years. A differential spectrum subtracting the normalized peak intensities of the NPDW from the averaged normalized peak intensities of the fjord sample (Fig. 7D) reveals the long-term modifications on the molecular level. Negative remainders indicate peaks with higher relative intensities in NPDW DOM (Fig. 7C), while positive remainders indicate peaks with higher relative intensities in Svalbard fjord DOM (Fig. 7B). Compounds that were relatively more abundant in the Svalbard fjord compared to the NPDW widely spread over the van Krevelen space (O/C versus H/C ratios). Characteristic NPDW molecules, on the other hand, occupy distinct regions in the van Krevelen diagram with most molecules being confined to the center and a few in the condensed aromatic carbon region at low O/C and H/C ratios. However, the average O/C (Svalbard DOM 0.49±0.08, NPDW DOM 0.39±0.18) and H/C (Svalbard DOM 1.25±0.14, NPDW DOM 1.18±0.36) ratios of the characteristic molecules are still similar for both environments, though the ranges are more variable. The difference between DOM from the Arctic and Pacific becomes more obvious if the average molecular mass of only those molecules that varied between the sites is compared:

Compounds that had a higher abundance in the fjords had an average molecular mass of 357, those that were higher in NPDW 468 Da. Presumably, the smaller molecules represent less refractory DOM and are selectively removed during the transport and transformation from the surface into the deep ocean. The DBC molecules were enriched in the NPDW presumably because black carbon is barely modified during the aging process in the deep sea (Dittmar and Paeng 2009) and due to the removal of other molecules, DBC increases in relative abundance.

Conclusions. Here we could show that the resident, highly active microbial community in Arctic fjords rapidly decomposed DOM that was produced during phytoplankton blooms in spring and early summer. The phytoplankton bloom did not leave any detectable imprint neither on the molecular composition of DOM, nor on its bulk concentration. This is most remarkable given the fact that we considered for our molecular characterization a total of 11,630 different compounds that were distinguished by their molecular mass. Because Arctic waters are an important source of DOM to the deep ocean, this finding is of major relevance. Assuming that our findings are representative for high latitude marine systems in general, we conclude that autochthonous seasonal DOC production in plankton bloom situations possibly contributes only marginally to the bulk DOC pool in these regions, likely also including the major deep water formation sites. In the course of global ocean circulation, Arctic DOM will eventually decompose and undergo changes in its molecular composition. Although Arctic fjord DOM and aged, refractory deep Pacific Ocean water DOM shared 95% of the detected masses, the relative intensity distributions of the components reveal significant differences in the molecular DOM composition, reflecting ageing during deep ocean circulation.

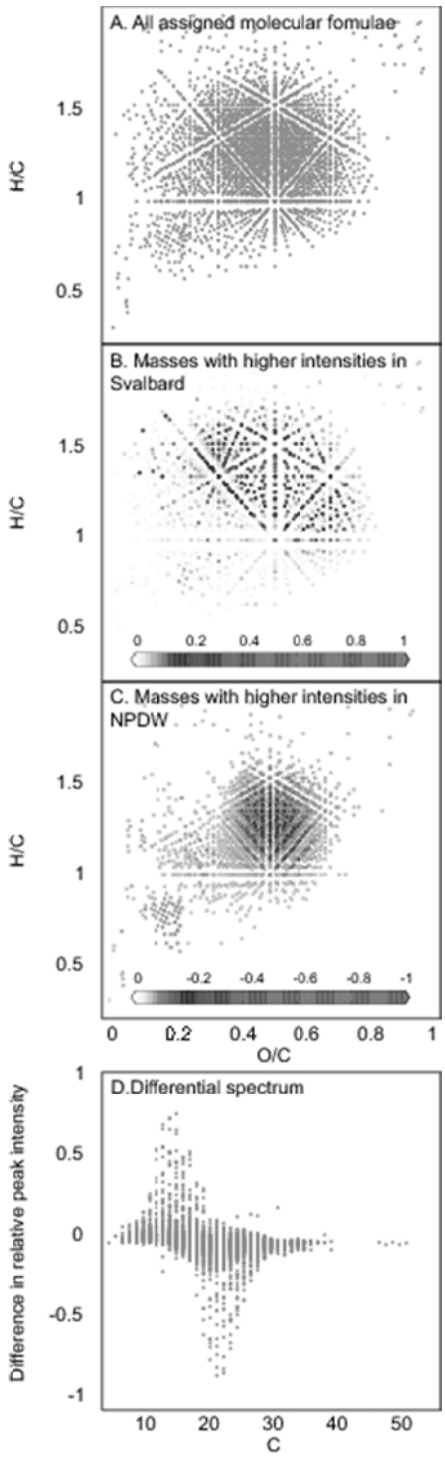

Fig. 7. Van Krevelen diagrams of A. all detected masses with assigned molecular formulae, B. molecules whose intensities were on average higher in Svalbard than in NPDW (n = 1973, the gray-scale bar denotes differences of relative FT-ICR-MS signal intensities), C. molecules whose intensities were on average higher in NPDW than in Svalbard (n = 2423), and D. average NPDW peak intensities subtracted from average Svalbard peak intensities against the number of C atoms per molecule.

2.7 Contributors

The study was initiated and designed by J.N. and T.D.. H.O. and J.N. carried out field work and sample preparation. H.O. performed all microbiological analyses and ultrahigh resolution mass spectrometry, J.N. did DBC analysis. All data analysis including statistical analysis was done by H.O., advised by J.N. and T.D.. All authors contributed to data interpretation and general discussion. H.O. wrote the manuscript with input from J.N. and T.D..

2.8 Acknowledgements

Captain Stig Hennigsen, the MS Farm Crew and the Ny Ålesund Marine Laboratory staff are thanked for help during the field campaign. The authors are most grateful to Katrin Klaproth for support in FT-ICR-MS analyses and to Matthias Friebe for DOC measurements. Gabriel A. Singer contributed helpful insights into multivariate statistics, Bert Engelen and Bernhard Fuchs kindly provided access to DGGE and Flow Cytometry, respectively, and valuable support for initial data interpretation. The very helpful comments of two anonymous reviewers greatly improved an earlier version of this manuscript. Financial support for this study came from Schlumberger Limited, the Max Planck Institute for Marine Microbiology, and the University of Oldenburg in Germany.

2.9 References

Amon, R. M. W., H. P. Fitznar, and R. Benner. 2001. Linkages among the bioreactivity, chemical composition, and diagenetic state of marine dissolved organic matter. Limnol. Oceanogr. **46:** 287-297.

Amon, R. M. W., G. Budeus, and B. Meon. 2003. Dissolved organic carbon distribution and origin in the Nordic Seas: Exchanges with the Arctic Ocean and the North Atlantic. J. Geophys. Res.-Oceans **108.**

Amon, R. M. W., and B. Meon. 2004. The biogeochemistry of dissolved organic matter and nutrients in two large Arctic estuaries and potential implications for our understanding of the Arctic Ocean system. Mar. Chem. **92:** 311-330.

Anderson, M. J. 2006. Distance-Based Tests for Homogeneity of Multivariate Dispersions. Biometrics **62:** 245-253.

Arnosti, C., and B. B. Jørgensen. 2003. High activity and low temperature optima of extracellular enzymes in Arctic sediments: implications for carbon cycling by heterotrophic microbial communities. Mar. Ecol.-Prog. Ser. **249:** 15-24.

Arnosti, C., and A.D. Steen. 2013. Patterns of extracellular enzyme activities and microbial metabolism in an Arctic fjord of Svalbard and in the northern Gulf of Mexico: contrasts in carbon processing by pelagic communities. Front. Microbiol. **4**: 318.

Azam, F. 1998. Microbial Control of Oceanic Carbon Flux: The Plot Thickens. Science **280:** 694-696.

Berman, T., and D. A. Bronk. 2003. Dissolved organic nitrogen: A dynamic participant in aquatic ecosystems. Aquat. Microb. Ecol. **31:** 279-305.

Bronk, D.A. 2002. Dynamics of DON. In: Biogeochemistry of marine dissolved organic matter. (eds. Hansell, D. A., and C. A. Carlson). Academic Press, 153-247.

Cottier, F., V. Tverberg, M. Inall, H. Svendsen, F. Nilsen, and C. Griffiths. 2005. Water mass modification in an Arctic fjord through cross-shelf exchange: The seasonal hydrography of Kongsfjorden, Svalbard. J. Geophys. Res.-Oceans **110:** 18.

Dittmar, T. 2008. The molecular level determination of black carbon in marine dissolved organic matter. Org. Geochem. **39:** 396-407.

Dittmar, T., and G. Kattner. 2003. The biogeochemistry of the river and shelf ecosystem of the Arctic Ocean: a review. Mar. Chem. **83:** 103-120.

Dittmar, T., B. Koch, N. Hertkorn, and G. Kattner. 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. Limnol. Oceanogr.-Methods **6:** 230-235.

Dittmar, T., and B. P. Koch. 2006. Thermogenic organic matter dissolved in the abyssal ocean. Mar. Chem. **102:** 208-217.

Dittmar, T., and J. Paeng. 2009. A heat-induced molecular signature in marine dissolved organic matter. Nature Geosci. **2:** 175-179.

Duursma, E. K. 1963. The production of dissolved organic matter in the sea, as related to the primary gross production of organic matter. Neth. J. Sea Res. **2:** 85-94.

Eilertsen, H. C., J. P. Taasen, and J. M. Wesiawski. 1989. Phytoplankton studies in the fjords of West Spitzbergen: physical environment and production in spring and summer. J. Plankton Res. **11:** 1245-1260.

Elvebakk, A. 1994. A Survey of Plant Associations and Alliances from Svalbard. J. Veg. Sci. **5:** 791-802.

Fandino, L., B. , L. Riemann, G. Steward, F. , R. A. Long, and F. Azam. 2001. Variations in bacterial community structure during a dinoflagellate bloom analyzed by DGGE and 16S rDNA sequencing. Aquat. Microb. Ecol. **23:** 119-130.

Flerus, R. and others 2012. A molecular perspective on the ageing of marine dissolved organic matter. Biogeosciences **9:** 1935-1955.

Grossart, H. P., F. Levold, M. Allgaier, M. Simon, and T. Brinkhoff. 2005. Marine diatom species harbour distinct bacterial communities. Environ Microbiol. 7(6):860-73.

Hansell, D. A. 2013. Recalcitrant Dissolved Organic Carbon Fractions. Annu. Rev. Mar. Sci **5**: 421-445.

Hedges, J. I. 1992. Global biogeochemical cycles - progress and problems. Mar. Chem. **39:** 67-93.

Hood, E., and L. Berner. 2009. Effects of changing glacial coverage on the physical and biogeochemical properties of coastal streams in southeastern Alaska. J. Geophys. Res., G, Biogeosci. **114:** 03001-03010.

Hop, H. and others 2002. The marine ecosystem of Kongsfjorden, Svalbard. Polar Res. **21:** 167-208.

Iversen, K. R., and L. Seuthe. 2011. Seasonal microbial processes in a high-latitude fjord (Kongsfjorden, Svalbard): I. Heterotrophic bacteria, picoplankton and nanoflagellates. Polar Biol. **34:** 731-749.

Jaffé, R. and others 2013. Global Charcoal Mobilization from Soils via Dissolution and Riverine Transport to the Oceans. Science **340:** 345-347.

Jiao, N. and others 2010. Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. Nat. Rev. Microbiol. **8:** 593-599.

Jiao, N. 2011. The microbial carbon pump and the oceanic recalcitrant dissolved organic matter pool. Nat. Rev. Microbiol. **9.**

Kim, S., L. A. Kaplan, and P. G. Hatcher. 2006. Biodegradable Dissolved Organic Matter in a Temperate and a Tropical Stream Determined from Ultra-High Resolution Mass Spectrometry. Limnol. Oceanogr. **51:** 1054-1063.

Kirchman, D. L., Y. Suzuki, C. Garside, and H. W. Ducklow. 1991. High turnover rates of dissolved organic carbon during a spring phytoplankton bloom. Nature **352:** 612-614.

Koch, B. P., and T. Dittmar. 2006. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. Rapid Commun. Mass Spectrom. **20:** 926-932.

Koch, B. P., T. Dittmar, M. Witt, and G. Kattner. 2007. Fundamentals of Molecular Formula Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter. Anal. Chem. **79:** 1758-1763.

Koch, B. P., M. R. Witt, R. Engbrodt, T. Dittmar, and G. Kattner. 2005. Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Geochimica Et Cosmochimica Acta **69:** 3299-3308.

Kujawinski, E. B. and others 2009. Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. Geochim. Cosmochim. **73:** 4384-4399.

Marie, D., F. Partensky, S. Jacquet, and D. Vaulot. 1997. Enumeration and Cell Cycle Analysis of Natural Populations of Marine Picoplankton by Flow

Cytometry Using the Nucleic Acid Stain SYBR Green I. Appl. Environ. Microbiol. **63(1):**186-93.

Mcclelland, J. W., R. M. Holmes, K. H. Dunton, and R. W. Macdonald. 2012. The Arctic Ocean Estuary. Estuaries and Coasts **35:** 353-368.

Muyzer, G., E. C. Dewaal, and A. G. Uitterlinden. 1993. Profiling of Complex Microbial-Populations by Denaturing Gradient Gel-Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes-Coding for 16s Ribosomal-RNA. Appl. Environ. Microbiol. **59:** 695-700.

Oksanen, J. and others 2012. vegan: Community Ecology Package. R package version 2.0-4. http://CRAN.R-project.org/package=vegan.

Opsahl, S., R. Benner, and R. M. W. Amon. 1999. Major flux of terrigenous dissolved organic matter through the Arctic Ocean. Limnol. Oceanogr. **44:** 2017-2023.

Pomeroy, L. R. 1974. Oceans food web, a changing paradigm. Bioscience **24:** 499-504.

Rudels, B., and D. Quadfasel. 1991. Convection and deep water formation in the Arctic Ocean-Greenland Sea System. J. Marine Syst. **2:** 435-450.

Singer, G. A. and others 2012. Biogeochemically diverse organic matter in Alpine glaciers and its downstream fate. Nature Geosci. **5:** 710-714.

Stenson, A. C., A. G. Marshall, and W. T. Cooper. 2003. Exact Masses and Chemical Formulas of Individual Suwannee River Fulvic Acids from Ultrahigh Resolution Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectra. Anal. Chem. **75:** 1275-1284.

Stevens, H., T. Brinkhoff, and M. Simon. 2005. Composition of free-living, aggregate-associated and sediment surface-associated bacterial communities in the German Wadden Sea. Aquat. Microb. Ecol. **38:** 15-30.

Stuiver, M., P. D. Quay, and H. G. Ostlund. 1983. Abyssal Water Carbon-14 Distribution and the Age of the World Oceans. Science **219:** 849-851.

Svendsen, H. and others 2002. The physical environment of Kongsfjorden-Krossfjorden, an Arctic fjord system in Svalbard. Polar Res. **21:** 133-166.

Teske, A., A. Durbin, K. Ziervogel, C. Cox, and C. Arnosti. 2011. Microbial Community Composition and Function in Permanently Cold Seawater and Sediments from an Arctic Fjord of Svalbard. Appl. Environ. Microbiol. **77:** 2008-2018.

Wängberg, S.-Å., K. I. M. Andreasson, K. Gustavson, T. Reinthaler, and P. Henriksen. 2008. UV-B effects on microplankton communities in Kongsfjord, Svalbard: A mesocosm experiment. J. Exp. Mar. Biol. Ecol. **365:** 156-163.

Wheeler, P. A. and others 1996. Active cycling of organic carbon in the central Arctic Ocean. Nature **380:** 697-699.

Wheeler, P. A., J. M. Watkins, and R. L. Hansing. 1997. Nutrients, organic carbon and organic nitrogen in the upper water column of the Arctic Ocean: implications for the sources of dissolved organic carbon. Deep-Sea Res. PT II: Topical Studies in Oceanography **44:** 1571-1592.

Williams, P. M., and E. R. M. Druffel. 1987. Radiocarbon in dissolved organic matter in the central North Pacific Ocean. Nature **330:** 246-248.

Zhou, J. Z., M. A. Bruns, and J. M. Tiedje. 1996. DNA recovery from soils of diverse composition. Appl. Environ. Microbiol. **62:** 316-322.

3 Do marine microorganisms really produce refractory dissolved organic matter?

Helena Osterholz[1], Jutta Niggemann[1], Helge-Ansgar Giebel[2], Meinhard Simon[2], Thorsten Dittmar[1]

manuscript in preparation

[1]Research Group for Marine Geochemistry (ICBM-MPI Bridging Group), [2]Biology of Geological Processes, Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky University Oldenburg, Carl-von-Ossietzky-Strasse 9-11, 26129 Oldenburg, Germany

3.1 Abstract

Dissolved organic matter (DOM) in the oceans constitutes a major carbon pool involved in global biogeochemical cycles[1]. The existence of this reservoir, containing almost 700 Gt of carbon, seems paradoxical considering its potential as nutrient source to microbial life and its occurrence in an aqueous, oxic environment. More than 96% of the marine DOM appear resistant to microbial degradation on timescales of several thousands of years[2] whereas microbial transformation of labile into refractory DOM (RDOM) appears to be fast[3-5]. This implies that RDOM is produced at enormously high efficiency, by far higher than needed to sustain the global RDOM pool. Challenging to any experimental approach is the involved millennium timescale. We overcame this issue by investigating the microbial formation and transformation of DOM on an unprecedented molecular level during a 3-years mesocosm study. We show that most of the apparent microbially produced RDOM is molecularly different from millennium-old oceanic RDOM. Only a tiny fraction (<3%) of the net community production was channeled into a form of RDOM molecularly undistinguishable from DOM in the deep ocean, consistent with global estimates on the production, turnover and accumulation of DOM in the global ocean.

3.2 Results & Discussion

A useful concept in marine biogeochemistry is the operational definition of DOM reactivity fractions[2]. Labile DOM comprises organic compounds with lifetimes from minutes to days, while semi-labile and semi-refractory DOM are turned over within 1 to 20 years. RDOM is the most ubiquitous fraction and resides in the world's oceans for thousands of years. It may be formed through abiotic processes[6-9], but biological production of recalcitrant DOM has also been proposed[10]. The microbial production of DOM from simple substrates that persists further degradation for months[3,5,11] to years[4] was demonstrated based on quantification of bulk dissolved organic carbon (DOC) and its fraction of amino acids and carbohydrates. However, less than 5% of deep ocean RDOM consists of these molecularly defined

compounds classes[12]. An inherent property of RDOM is its enormous molecular diversity that can only be revealed through ultrahigh-resolution analytical techniques, most prominently Fourier-transform ion cyclotron resonance mass spectrometry (FT ICR MS)[13]. The molecular diversity of microbially formed DOM assessed by this approach is very similar to natural oceanic RDOM[14]. A currently unresolved conundrum is the observation that RDOM is apparently produced by microorganisms from labile substrates with far higher efficiency than required to sustain the characteristics of the natural DOM pool[14]. Hence, the oceanic DOM pool would be several times larger if the experimentally derived production rates were representative[14]. This obvious paradox with respect to RDOM formation is of concern considering the fact that minor changes in the global RDOM pool could substantially influence atmospheric carbon dioxide concentration and the radiation balance on Earth.

A major caveat of all laboratory experiments is their short duration, compared to the millennium-scale stability of oceanic RDOM. The question remains whether the "non-labile" DOM detected in experimental studies is truly equal to oceanic RDOM. To overcome this issue we here define RDOM not only by its stability on an experimental time scale (years), but also by its molecular composition. We define "truly" refractory DOM as equal in molecular composition to abyssal RDOM in terms of relative abundance of individual DOM constituents. For this assessment we combined conventional molecular analysis specifically targeting amino acids and carbohydrates with non-targeted ultrahigh-resolution FT-ICR MS. Thousands of molecular formulae can be identified in DOM via this technique[13,15], and further fragmentation studies provide detailed structural information beyond this level[16]. We hypothesized that RDOM, as found in the deep ocean, is generated by diverse microbial communities, but at much lower efficiency than previously assumed based on experimental approaches.

To test our hypothesis, we inoculated low-DOC artificial seawater (~7 μmol C L^{-1}) containing all essential inorganic nutrients with a coastal North Sea phytoplankton and bacterial community. This approach was chosen to include as many of the

naturally occurring interactions between primary producers and heterotrophic microbes as possible. Triplicate mesocosms (M1-M3) were incubated at a 12:12 hour light-dark cycle and incubated at room temperature (~22°C). After 167 days, a portion of each mesocosm was filtered (1.2 µm) to exclude large particles and further incubated in the dark (M1-M3*dark*). The mesocosms were subsampled at increasing intervals ranging from 2 to 300 days for a total of 1011 days. Over the duration of the experiment, algal and bacterial abundance and community composition, inorganic nutrients (NO_3^-, NO_2^-, NH_4^+, PO_4^{3-}), total hydrolysable dissolved amino acids (THDAA), carbohydrates, DOC and the molecular composition of solid-phase extracted DOM were monitored. North Pacific Deep Water (NPDW), one of the oldest water masses on Earth, served as reference material for the composition of refractory DOM[17].

Three consecutive phases of the mesocosm experiment were identified: phytoplankton blooms, the post-bloom stage and the net-heterotrophic phase of DOM consumption.

I. Phytoplankton blooms. Between day 1 and 56, favorable experimental conditions with high nutrient supply resembling conditions in upwelling regions or coastal seas[18], induced a sequence of phytoplankton blooms: the first one was dominated by diatoms, followed by blooms with high proportions of flagellates and prymnesiophytes with maxima at days 21 and 38, similar to the natural phytoplankton succession[19] (Fig. 1, Table S2.1). Nitrate and ammonium were depleted below the detection limit within one week, while phosphate was never limiting. Bacteria reached 0.2-1×10^7 cells mL^{-1} (Fig. 1) and diverse communities (Bray-Curtis dissimilarity 0.64±0.1, Fig. 2b), dominated by members of the *Rhodobacteraceae* family and some *Bacteroidetes* representatives, developed (Tab. S2.2). About 100 µmol DOC L^{-1} were generated during the bloom phase regardless of phytoplankton community composition (Fig. 3). THDAA as percent of DOC peaked during the first phytoplankton bloom with up to 10.5% (Fig. S2.4e), indicating freshly produced labile DOM[20]. Also the relative proportion of individual amino acids was consistent with fresh DOM according to the calculated degradation

state index[21] (Fig. S2.4g). Molecular analysis of solid-phase extracted DOM[22] by FT ICR MS revealed that the number of resolved molecular formulae rapidly increased during the first days of the incubation to levels of deep-sea DOM (5800 detected formulae in NPDW). Notably, an increasing proportion of the new molecular formulae was identical to those found in the deep ocean (NPDW reference, up to 50% for routine analyses, Fig. S2.4o). With a more detailed investigation of mesocosm DOM via FT ICR MS at increased instrument sensitivity settings the agreement of molecular formulae could be raised to 82-97%. Based on the presence of molecular formulae, the experimental DOM was almost undistinguishable from deep-sea DOM within less than two months of incubation. This is most remarkable, because among the almost 8000 molecular formulae identified (with the same instrumentation) in the exometabolome of a pure *Pseudovibrio* strain less than 2% were identical to deep-sea DOM (NPDW)[23]. We attribute the high number of deep-sea molecular formulae in our experiment to the high microbial diversity in the mesocosms.

II. Post-bloom phase. Low chlorophyll *a* fluorescence prevailed in mesocosms M1 and M3 after day 56. In M2, a constant increase in fluorescence until day 785 was related to N_2-fixation in that particular mesocosm. The AA-C/DOC ratio fell below 3% in all light and dark mesocosms and remained low until day 1011; also the amino acid degradation state was consistent with aged DOM (Fig. S2.4g). Between days 167 and 370, most algal and many bacterial cells lysed, which led, together with recycled production, to DOC concentrations of up to 350 µmol C L^{-1} in the light mesocosms (Fig. 3). The dark mesocosms that started at day 167 showed little DOC production, presumably derived from decaying cells.

III. Net-heterotrophic phase. From day 370, consumption processes prevailed over new DOM production: bacterial numbers remained low and DOC concentrations decreased by ~200 µmol C L^{-1} in M1 and M3, but increased by 360 µmol C L^{-1} in M2. In the dark mesocosms, DOC concentrations decreased between 41 and 58%. Despite strong quantitative changes in DOC concentration, molecular DOM

composition changed very little (Fig. 2a, Bray Curtis dissimilarity after day 300: 0.27±0.16, supplementary material).

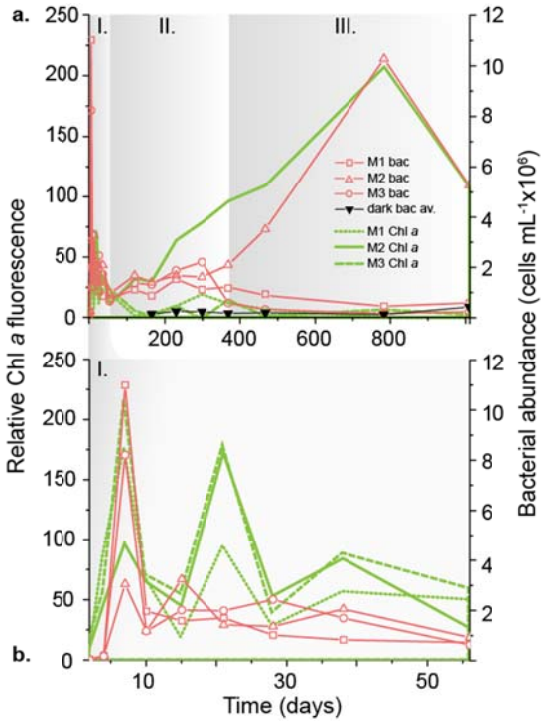

Figure 1: Phytoplankton blooms and bacterial abundance. Relative Chlorophyll *a* fluorescence (M1-M3 Chl *a*) and bacterial abundance of the light-dark (M1-M3 bac) and dark (dark bac av: mean bacterial abundance of triplicates) mesocosm over the course of the experiment (a) and close up of the three phytoplankton blooms until day 56 (b). The standard deviation of triplicates (dark bac av) is within size of symbols.

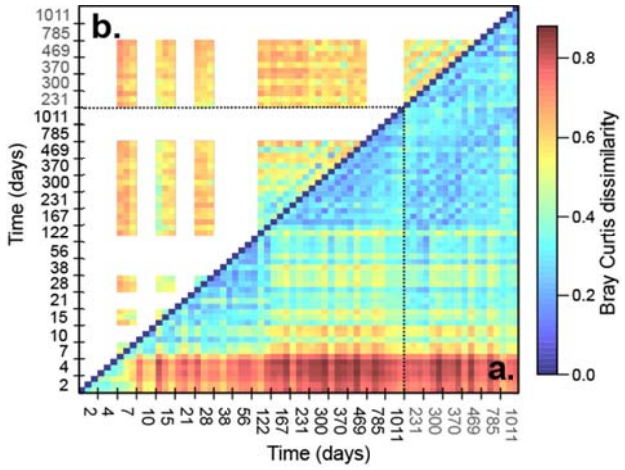

Figure 2: Dissimilarity analysis of molecular DOM and microbial community composition. Bray Curtis dissimilarity of samples based on the relative intensities of 9264 molecular formulae (a, lower triangle) and microbial community composition assessed via denaturing gradient gel electrophoresis fingerprinting (b, upper triangle, data not available for all sampling points) over time. Light (days 2-1011) and dark (days 231-1011) mesocosms are separated by the dotted line. Red color denotes a high dissimilarity between samples while blue color denotes more similar samples.

The diverse phytoplankton and bacterial community in our batch incubations produced molecularly highly diverse DOM with an apparently enormous similarity to deep-sea RDOM. The DOM production in the copious amount of >70 µmol DOC L^{-1} remaining after three years, after initial injection of 40 µmol inorganic nitrogen from ammonium and nitrate as the limiting nutrients, largely exceeds the amount of RDOM needed to sustain the global inventory[2]. In this sense, our study is consistent with the above mentioned conundrum revealed by previous studies[4,5,14]. If, however, the molecular similarity with deep-sea RDOM is extended towards the level of molecular structures and abundance patterns, a very different picture appears. One molecular formula may represent a multitude of structural isomers. To reveal structural features behind molecular formulae we fragmented a total of 32 abundant formulae of deep-sea DOM (NPDW) and a pooled M3 sample in the FT ICR MS (Fig. 4). The major fragmentation patterns revealed the prevalence of organic acids in both samples (i.e., combinations of CO_2, H_2O and CH_4O fragments); other structural features such as esters, ethers, or peptides were

rare in both samples. Despite this overall structural similarity, the abundance patterns of the respective fragments were very different (Fig. 4), indicating that the structural composition of deep-sea DOM and DOM produced in the mesocosms was not identical.

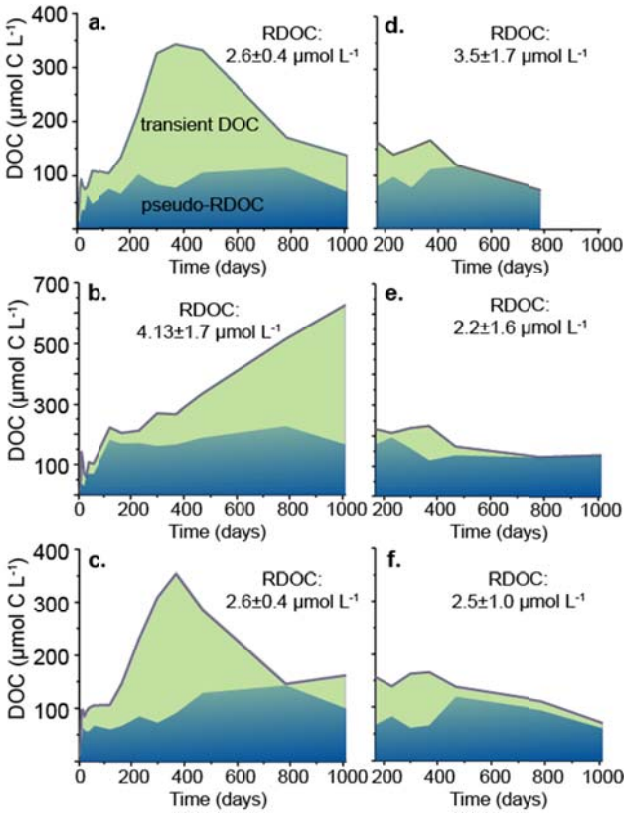

Figure 3: DOM fractions according to chemical composition and reactivity. DOC concentration and contributions of pseudo-RDOC (blue) and transient DOC (green) in light (a-c, M1-M3) and dark (d-f, M1-M3*dark*) mesocosms. RDOC concentration was averaged for samples taken after day 56. The last sample for M1*dark*, day 1011 (d) was lost during sample processing.

Based on general chemical properties and molecular characteristics, we defined three fractions of microbially produced DOM (Tab. 1, Fig. 3). The "transient DOM" constituted the fraction most dissimilar to RDOM and was also most dynamic. It slowly accumulated but was then almost completely degraded in the net heterotrophic phase. The high C/N ratio (Fig. S4c) suggested a significant contribution of carbohydrates which could be confirmed by the quantification of

dissolved combined carbohydrates (DCCHO) at day 1011 (M1, M3 ~3 µmol L^{-1}, M2 ~26 µmol L^{-1}, M1-M3*dark* <1 µmol L^{-1}, Tab. S2.3). We defined "pseudo-RDOM" as being very similar but not fully identical to deep-sea DOM (NPDW), both with respect to chemical properties and molecular characteristics. Pseudo-RDOM was formed during the bloom phase and remained at concentrations of ~100 µmol C L^{-1}.

We propose that much of the DOM in this class had previously been termed RDOM because of its longevity on experimental timescales and chemical similarity to RDOM. Ultrahigh-resolution mass spectrometry now enabled us to define the smallest DOM fraction: it is equal to deep-sea DOM (NPDW) in terms of presence and relative intensities of resolved masses. To match this definition, the molecular formulae detected in deep-sea DOM cannot be larger in relative intensity to their respective counterparts in the mesocosm DOM sample, on a given confidence level (see S1.6 for details). Based on this definition, at most 2.9±1.4 µmol L^{-1} refractory DOM existed in the mesocosm DOM at day 56 and thereafter, representing <0.18 to 0.36% of the net community production in the mesocosms (NCP; new production based on inorganic nitrogen and phosphate uptake).

Our estimates are well within the same order of magnitude of global RDOM production. Global estimates report a NCP of 0.6 Pmol C yr^{-1}, 0.63% of which are channeled into the RDOM pool[2,24]. Also the observed production of "pseudo-RDOM" is consistent with estimates on the global production of semi-labile and semi-refractory DOM[2,24]. Photochemical alteration[8], thermogenic processes[9] and other abiotic factors may be significant for the formation of RDOM, but our study indicates that diverse microbial communities produce DOM that is similar in reactivity and composition to RDOM found in the global ocean. We were able to shed light on the production of large amounts of seemingly refractory DOM in mesocosm experiments and conclude that less RDOM than previously thought[10], but still enough to sustain the global RDOM pool, is produced by microorganisms. The underlying mechanisms behind the relationship between NCP and RDOM production remain to be understood, particularly in consideration of the yet unknown implications of global change[25] on DOM dynamics.

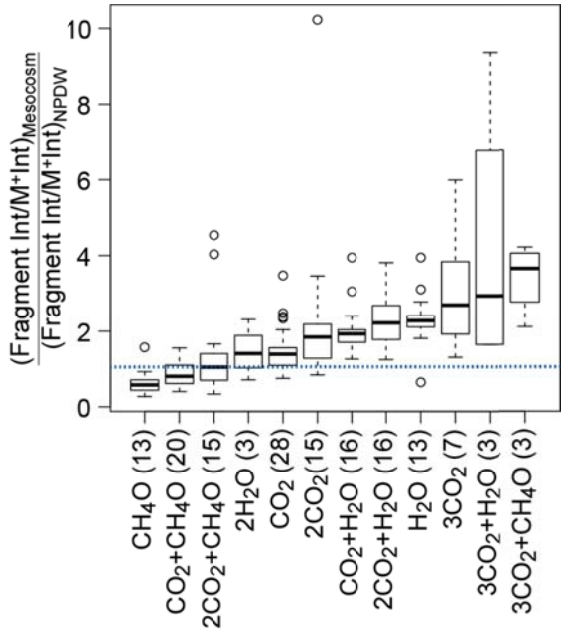

Figure 4: Structural diversity revealed by fragmentation. FT ICR MS collision fragmentation was conducted for six nominal masses contained in NPDW and mesocosm DOM. The fragmentation patterns describe the molecular dissimilarity on structural level. The boxplot depicts the ratio of the fragment intensity relative to/divided by the respective parent ion in mesocosm DOM to the fragment intensity relative to/divided by the parent ion intensity in NPDW DOM by fragment type. The same structure behind a molecular formula would yield a ratio of 1 (dotted blue line). The number behind the fragment molecular formula denotes the number of times the fragment was detected. The horizontal line in each box is the median, the hinges enclose 1st to 3rd quartile, and the notches represent upper and lower adjacent values.

Table 1: Reactivity of DOM fractions. Three fractions of DOM that are produced during the mesocosm experiment, assigned by their chemical properties and including proposed reactivity and formation processes.

Name	Chemical definition	Mathematical definition	Observed reactivity	Observed formation
RDOM (deep sea DOM)	- Identical to deep sea DOM in terms of bulk parameters (C/N ratio, solid-phase extractability[22] (2/3 of DOC is solid-phase extractable, i.e., SPE-DOC), amino acid content and composition) and molecular parameters (presence/absence and intensity distribution of molecular formulae)	[RDOC] = %NPDW × [SPE-DOC] × 3/2	semi-refractory to refractory (stable for thousands of years)	-produced in all mesocosms mainly during phase of new production
pseudo-RDOM	- Identical to deep sea DOM in terms of bulk parameters (C/N ratio, solid-phase extractability (2/3), amino acid content and composition) - Similar to deep sea DOM in terms of presence/absence of molecular formulae - Different to deep sea DOM in terms of intensity distribution of molecular formulae and molecular structures	[pseudo-RDOC] = [SPE-DOC] × 3/2 – [RDOC]	semi-refractory to refractory (stable for years)	produced in all mesocosms mainly during phase of new production
transient DOM	- Non-extractable via solid phase extraction from seawater - Contains little to no nitrogen - Contains at least ~1/3 carbohydrates	[transient-DOC] = [DOC] – [RDOC] – [pseudo-RDOC]	Semi-labile (half-life of months)	Accumulation during first phase of recycled production

3.3 Methods summary

9.9 L of artificial seawater (marine broth[26] without yeast extract and peptone, DOC 7.4 µmol C L^{-1}) were inoculated with 0.1 L of coastal North Sea water (May 30 2011, N53°31'48" E7°9'26", prefiltered 200 µm) as a source of microorganisms. Initial concentrations of inorganic nutrients were 20 µmol L^{-1} nitrate, 20 µmol L^{-1} ammonium, 25 µmol L^{-1} phosphate and 16 µmol L^{-1} silicate. Silicate concentrations over time were not reported since the incubations were performed in glass vessels. The incubation was performed in triplicate at room temperature (av. 22.5°C), in a 12:12 h light:dark cycle (light source: 120 µmol m^{-2} s^{-1}, 400-700 nm) and constant stirring. Controls consisted of ultrapure water to monitor external contamination; the

DOC concentration in the controls never exceeded 12 µmol C L^{-1}. Subsamples were taken in irregular intervals to monitor bacterial and microalgal community, nutrient availability as well as DOM quantity and quality. After 167 days, a proportion of the incubation medium was filtered through 1.2 µm to remove particles and larger organisms and subsequently kept in the dark to exclude primary production (M1-M3*dark*). They were sampled identically to the continued light mesocosms. The microalgal community composition was determined via direct counts under an inverted microscope at five timepoints. At all other sampling points, chlorophyll *a* fluorescence was monitored due to the limited volume. Microbial cell numbers were enumerated via flow cytometry[27] (Accuri C6, BD Biosciences) and the community composition was assessed via DGGE (denaturing gradient gel electrophoresis) fingerprinting of 16S rRNA gene fragments[28] and sequencing of excised bands. Inorganic nutrients (NH_4^+, NO_3^-, NO_2^-, PO_4^{3-}) were analyzed using standard photochemical methods[29,30]. DOC and TDN concentration were quantified by high temperature catalytic oxidation[31] (Shimadzu TOC-VCPH/CPN Total Organic Carbon Analyzer, ASI-V autosampler, TNM-1 module). Total hydrolysable dissolved amino acids (THDAA) were quantified by HPLC (Agilent Technologies, LC 1200) after acid hydrolysis[32]. Solid phase extraction[22] was used to concentrate and desalt the samples for the characterization of the molecular DOM composition via 15 T Solarix ESI FT ICR MS (Bruker Daltonics) in negative mode. Methanol extracts were diluted to 20 mg C L^{-1} at a ratio of 1:1 (v/v) MeOH to water, spiked with an internal standard (octaethylene glycol) and 500 scans were obtained with an ion accumulation time of 0.25 s. Mass spectra were exported and processed with in-house Matlab routines for peak matching and molecular formula assignment[15], yielding 11365 molecular formula assignments for all samples including the 70 mesocosm samples as well as repeat measurements of a North Pacific Deep Water (NPDW) reference sample[17] and the solid phase extracted North Sea inoculum (NSW). The molecular DOM composition of the mesocosm samples and the references was compared on the basis of normalized peak intensities of detected masses with assigned molecular formulae. Fragmentation experiments for 3 masses of 2 homologous series were conducted to explore the structural resemblance of

molecular formulae assigned to NPDW DOM and a pooled mesocosm DOM sample (M3, days 370-785).

3.4 Author contributions

TD, JN, HO and MS conceived the study, HO performed all analyses and data handling. HAG performed flow cytometric analysis and enumeration of microorganisms. All authors contributed to data interpretation. HO wrote the manuscript with significant input from all coauthors.

3.5 Acknowledgements

The authors thank Katrin Klaproth for technical assistance with FT ICR MS analyses, Matthias Friebe for DOC and TDN analysis, Birgit Kuerzel for amino acid analysis and Heike Simon for carbohydrate analyses. Dr. Christoph Feenders contributed to data interpretation. Funding was provided by the Max Planck Institute for Marine Microbiology (Bremen, Germany), the University of Oldenburg (Germany), and DFG (TRR 51).

3.6 References

1 Hedges, J. I. Global biogeochemical cycles - progress and problems. *Mar. Chem.* **39**, 67-93 (1992).

2 Hansell, D. A. Recalcitrant Dissolved Organic Carbon Fractions. *Annual Review of Marine Science* **5**, 421-445 (2013).

3 Tranvik, L. J. Microbial transformation of labile dissolved organic matter into humic-like matter in seawater. *Fems Microbiology Ecology* **12**, 177-183 (1993).

4 Ogawa, H., Amagai, Y., Koike, I., Kaiser, K. & Benner, R. Production of Refractory Dissolved Organic Matter by Bacteria. *Science* **292**, 917-920, (2001).

5 Brophy, J. E. & Carlson, D. J. Production of biologically refractory dissolved organic carbon by natural seawater microbial populations. *Deep Sea Research Part A. Oceanographic Research Papers* **36**, 497-507 (1989).

6 Harvey, G. R., Boran, D. A., Chesal, L. A. & Tokar, J. M. The structure of marine fulvic and humic acids. *Mar. Chem.* **12**, 119-132, (1983).

7 Keil, R. G. & Kirchman, D. L. Abiotic transformation of labile protein to refractory protein in seawater. *Mar. Chem.* **45**, 187-196, (1994).

8 Stubbins, A. *et al.* Illuminated darkness: Molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry. *Limnology and Oceanography* **55**, 1467-1477 (2010).

9 Dittmar, T. & Paeng, J. A heat-induced molecular signature in marine dissolved organic matter. *Nature Geoscience* **2**, 175-179 (2009).

10 Jiao, N. *et al.* Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nat Rev Microbiol* **8**, 593-599 (2010).

11 Gruber, D. F., Simjouw, J.-P., Seitzinger, S. P. & Taghon, G. L. Dynamics and Characterization of Refractory Dissolved Organic Matter Produced by a Pure Bacterial Culture in an Experimental Predator-Prey System. *Appl. Environ. Microbiol.* **72**, 4184-4191 (2006).

12 Dittmar, T. & Stubbins, A. in *Treatise on Geochemistry (Second Edition)* (eds Heinrich D. Holland & Karl K. Turekian) 125-156 (Elsevier, 2014).

13 Koch, B. P., Witt, M. R., Engbrodt, R., Dittmar, T. & Kattner, G. Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochimica Et Cosmochimica Acta* **69**, 3299-3308 (2005).

14 Koch, B. P., Kattner, G., Witt, M. & Passow, U. Molecular insights into the microbial formation of marine dissolved organic matter: recalcitrant or labile? *Biogeosciences Discuss.* **11**, 3065-3111 (2014).

15 Osterholz, H., Dittmar, T. & Niggemann, J. Molecular evidence for rapid dissolved organic matter turnover in Arctic fjords. *Mar. Chem.* **160**, 1-10 (2014).

16 Witt, M., Fuchser, J. & Koch, B. P. Fragmentation Studies of Fulvic Acids Using Collision Induced Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Analytical Chemistry* **81**, 2688-2694 (2009).

17 Green, N. W. *et al.* An intercomparison of three methods for the large-scale isolation of oceanic dissolved organic matter. *Mar. Chem.* (2014).

18 Dugdale, R. & Wilkerson, F. in *Primary Productivity and Biogeochemical Cycles in the Sea* Vol. 43 *Environmental Science Research* (eds P.G. Falkowski, A.D. Woodhead, & K. Vivirito) Ch. 7, 107-122 (Springer US, 1992).

19 Taylor, A. H., Harbour, D. S., Harris, R. P., Burkill, P. H. & Edwards, E. S. Seasonal succession in the pelagic ecosystem of the North Atlantic and the utilization of nitrogen. *Journal of Plankton Research* **15**, 875-891 (1993).

20 Benner, R. in *Biogeochemistry of Marine Dissolved Organic Matter* (eds D. A. Hansell & C. A. Carlson) Ch. 3, 59-90 (Academic Press, 2002).

21 Dauwe, B. & Middelburg, J. J. Amino acids and hexosamines as indicators of organic matter degradation state in North Sea sediments. *Limnology and Oceanography* **43**, 782-798 (1998).

22 Dittmar, T., Koch, B., Hertkorn, N. & Kattner, G. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnology and Oceanography-Methods* **6**, 230-235 (2008).

23 Romano, S. *et al.* Exo-Metabolome of *Pseudovibrio* sp. FO-BEG1 Analyzed by Ultra-High Resolution Mass Spectrometry and the Effect of Phosphate Limitation. *PLoS ONE* **9**, e96038 (2014).

24 Hansell, D. A. & Carlson, C. A. Net community production of dissolved organic carbon. *Global Biogeochemical Cycles* **12**, 443-453 (1998).

25 Gruber, N. Warming up, turning sour, losing breath: ocean biogeochemistry under global change. *Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci.* **369**, 1980-1996 (2011).

26 Zobell, C. E. Studies on marine bacteria. I. The cultural requirements of heterotrophic aerobes. *Journal of Marine Research* **4**, 42-75 (1941).

27 Gasol, J. M. & Del Giorgio, P. A. Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. *Scientia Marina* **64**, 197-224 (2000).

28 Muyzer, G., Dewaal, E. C. & Uitterlinden, A. G. Profiling of Complex Microbial-Populations by Denaturing Gradient Gel-Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes-Coding for 16s Ribosomal-RNA. *Appl. Environ. Microbiol.* **59**, 695-700 (1993).

29 Schnetger, B. & Lehners, C. Determination of nitrate plus nitrite in small volume marine water samples using vanadium(III)chloride as a reduction agent. *Mar. Chem.* (2014).

30 Laskov, C., Herzog, C., Lewandowski, J. & Hupfer, M. Miniaturized photometrical methods for the rapid analysis of phosphate, ammonium, ferrous iron, and sulfate in pore water of freshwater sediments. *Limnology and Oceanography-Methods* **5**, 63-71 (2007).

31 Sugimura, Y. & Suzuki, Y. A high-temperature catalytic oxidation method for the determination of non-volatile dissolved organic carbon in seawater by direct injection of a liquid sample. *Mar. Chem.* **24**, 105-131 (1988).

32 Lindroth, P. & Mopper, K. High performance liquid chromatographic determination of subpicomole amounts of amino acids by precolumn fluorescence derivatization with o-phthaldialdehyde. *Analytical Chemistry* **51**, 1667-1674 (1979).

S1 Supplementary Methods

S1.1 Mesocosm set-up

At the start of the mesocosm experiment, 9.9 L artificial seawater (Difco MB 2216[1] without yeast extract and peptone, DOC 7.4 µmol C L^{-1}) were transferred into carefully acid-rinsed 10 L glass bottles and mixed with 0.1 L coastal North Sea water (Norderney, Germany, May 30 2011, 30 psu, 247 µmol C L^{-1}, pre-filtered 200 µm) in triplicates (M1, M2, M3). The mesocosms were incubated at room temperature (av. 22.5°C), 12:12 hour light:dark (120 µmol m^{-2} s^{-1}, 400-700 nm) and the water was constantly stirred. After 167 days, 1.5 L of each mesocosm were filtered through 1.2 µm glass fiber filters (Whatmann, precombusted 400°C, 4 hrs) to remove aggregates and phytoplankton, and then further incubated in the dark (M1*dark*, M2*dark*, M3*dark*). At each of the 17 sampling points, 300-400 mL were withdrawn from the 10 L bottles and further processed and preserved for the analyses.

S1.2 Phytoplankton and chlorophyll *a*

Relative chlorophyll *a* fluorescence was determined fluorometrically at every sampling interval (excitation 540 nm, emission 684 nm, TD-700 Fluorometer, Turner Designs, USA).

Of the inoculum (North Sea water NSW) and at five time points during the experiment (days 4, 21, 38, 231, 469), 50-100 mL of the mesocosms/inoculum were fixed with acidic Lugol's iodine solution (1 mL per 50 mL sample) and stored in the dark at 7°C. Enumeration and classification of phytoplankton cells was performed in triplicates by AquaEcology (Oldenburg, Germany, Table S2.1).

Additional confirmation of the presence of cyanobacteria that were not identified by direct counts could be obtained via flow cytometric investigation of the autofluorescence of live samples on day 473. High phycobilin fluorescence provides strong evidence for a population of *Synechococcus* sp. (1.6×10^5 cells mL^{-1}) in M2,

while 4.4 $\times 10^3$ cells mL^{-1} of *Prochlorococcus* sp. were found in M1. M3 did not exhibit phycobilin fluorescence on day 473.

S1.3 Microbial cell counts and community analysis

The determination of cell numbers of the free-living microorganisms was conducted with a BD Accuri C6 cytometer (BD Biosciences, USA) using SybrGreen I (Invitrogen™, United Kingdom) staining and the internal fluidics calibration of the device. The fixed samples (2% f.c. glutardialdehyde, Carl Roth, Germany) were thawed and filtered through 50 µm filters (Cell Trics, Partec) to remove larger particles. Volume verification was done using TruCount beads (BD) as described previously[2,3]. Data were processed by BD Accuri C6 C-Flow software (Version 1.026421).

For the analysis of the bacterial community composition by denaturing gradient gel electrophoresis (DGGE)[4] of PCR-amplified 16S rRNA gene fragments, 30 to 40 mL of the sample were passed through 0.2 µm PC filters (Millipore, USA). The filters were stored frozen at -20°C. DNA extraction and DGGE was conducted as described in Osterholz et al. (2014)[5].

Prominent bands were excised with a sterile scalpel, transferred into 49 µL of PCR-grade water and stored at -20°C. The DNA was reamplified and purified (Gold DNA purification kit, Quiagen, Netherlands) prior to sequencing at GATC Biotech AG (Germany). The obtained sequences were compared with those in GenBank using the Blast tool (www.ncbi.nlm.nih.gov) and results are summarized in Table S2.2. Sequences are deposited at GenBank under accession number KJ914554 to KJ914571.

S1.4 Inorganic nutrients

Subsamples for nutrient analysis (NH_4^+, NO_3^-, NO_2^-, PO_4^{3-}) were filtered through 0.2 µm polycarbonate filters (Isopore™ Membrane Filter, EMD Millipore Corporation, USA), preserved with $HgCl_2$ and stored at 7°C until analysis. PO_4^{3-} (limit of detection 0.8 µmol L^{-1}, limit of quantification 2.2 µmol L^{-1}) and NH_4^+ (limit of detection 1.2 µmol L^{-1}, limit of quantification 3.7 µmol L^{-1}) were analyzed via

micro-photometric methods[6]. NO_3^- (NO_x limit of detection 0.4 µmol L^{-1}, limit of quantification 1.0 µmol L^{-1}) and NO_2^- (limit of detection 0.2 µmol L^{-1}, limit of quantification 0.6 µmol L^{-1}) were determined via vanadium(III)chloride reduction[7].

S1.5 DOM quantification and characterization

Amino acid quantification. Aliquots for the quantification of total dissolved hydrolysable amino acids (THDAA) were filtered through GHP syringe filters (0.2 µm, Acrodisc, PALL Life Science, USA) and stored at -20°C. The amino acids were quantified after orthophthaldialdehyde precolumn derivatization as described in Lunau et al. (2006)[8] on a high performance liquid chromatography system (HPLC, 1200 series, Zorbax Eclipse XDB-C18-4.6x12.5 mm precolumn, Zorbax Eclipse XDB-C18-4.6x150 mm main column, Agilent Technologies, USA). Aspartic acid, glutamic acid, histidine, serine, arginine, glycine, threonine, alanine, tyrosine, methionine, valine, and phenylalanine were resolved and quantified via external standards (Fig. S2.1d-f). The degradation state index DI was calculated after Dauwe & Middelburg (1998, Fig. S2.1g)[9].

Carbohydrate quantification. Samples for dissolved total neutral carbohydrates (DCCHO) were prepared in the same way as for THDAA quantification, but only analyzed in the sample of day 1011. DCCHO (Tab. S2.3, glucose, mannose, galactose, fucose, rhamnose, arabinose, xylose) quantification was performed by HPLC with a Carbopac PA 1 column (Dionex, Thermo Fisher Scientific Inc., USA) and pulsed amperometric detection[10].

DOC/TDN quantification. Between 100 and 250 mL of sample were sequentially filtered through a 0.7 µm glass fiber (Whatman, United Kingdom, combusted 400°C, 4 h) and a 0.2 µm polycarbonate filter (Isopore™ Membrane Filter, EMD Millipore Corporation, USA) and acidified to pH 2 (HCl 25% p.a., Carl Roth, Germany). An aliquot of this acidified 0.2 µm filtrate was sampled for duplicate quantification of dissolved organic carbon (DOC, Fig. S2.1a) and total dissolved nitrogen (TDN, Fig. S2.1b) via high temperature catalytic combustion using a Shimadzu TOC-VCPH/CPN Total Organic Carbon Analyzer equipped with an ASI-V autosampler and a TNM-1 module[5]. For the calculation of C/N ratios, the

inorganic nitrogen concentration was subtracted from TDN (Fig. S2.1c). The error of DOC and TDN analysis was on average 4 % and 10 %, respectively. The DOC concentration in the control treatment (consisting of ultrapure water only) to monitor external DOC contamination, e.g. by use of solvents in the laboratory, did never exceed 12.3 µM.

Solid-phase extraction and FT ICR MS analysis. DOM was solid phase-extracted from the remaining filtered and acidified water (150-250 mL) with commercially available 100 mg modified styrene divinyl benzene polymer cartridges (PPL, Agilent, USA)[11]. After extraction, cartridges were rinsed with acidified ultrapure water (pH 2, HCl 25%, p.a., Carl Roth, Germany) to remove remaining salts, dried by flushing with nitrogen gas and eluted with 500 µl of methanol (HPLC-grade, Sigma-Aldrich, USA). Extracts were stored in amber vials at -20°C. The extraction efficiency varied between 14.6 and 85.2 % (average: 42.9±16.6 %) on carbon basis (Fig. S2.1h). The extractable DOC is calculated by multiplying the original DOC concentration in each sample by its determined extraction efficiency (Fig. S2.1i).

The mass spectra were obtained on a 15 Tesla Solarix FT ICR MS (Fourier Transform Ion Cyclotron Resonance Mass Spectrometer, Bruker Daltonics, USA) equipped with an electrospray ionization source (Bruker Apollo II) applied in negative mode. DOM extracts were diluted to a final DOC concentration of 20 mg C L^{-1} in a 1:1 (v/v) mixture of ultrapure water and methanol (HPLC-grade, Sigma-Aldrich, USA). For routine analyses, a total of 500 scans were accumulated per run in a mass window from 150 to 2000 Da. The spectra were calibrated with an internal calibration list using the Bruker Daltonics Data Analysis software package. The mass to charge, resolution and intensity were then exported and processed using in-house Matlab routines. Diluted octaethylene glycol solution (PEG-8, $H(OCH_2CH2)_8$-OH, ≥95% oligomer purity, Sigma Aldrich) was added to the diluted extracts as an internal standard before FT ICR MS analysis for better quantitative comparison of the spectra. The PEG-8 compound was kept at room temperature in water before use and thus underwent oxidation reactions. All samples were measured at the same time, and five peaks of the octaethylene glycol standard were

present in every mass spectrum and used for scaling of the samples. Molecular formulae were assigned to peaks with a minimum signal-to-noise ratio of 4 following the rules published in Koch et al. (2007)[12]. The number of assigned molecular formulae (Fig. S2.1j), relative peak intensity-weighted molecule mass (Fig. S2.1k), double bond equivalents (DBE, Fig. S2.1l), H/C ratio (Fig. S2.1m) and O/C ratio (Fig. S2.1n) were calculated for each sample.

Reference samples. The North Sea inoculum (NSW), was solid phase extracted as described for the mesocosm samples and analyzed repeatedly via FT ICR MS (n=15) to provide an estimate of the background DOM introduced into the mesocosms at the start.

The refractory DOM reference sample was obtained from North Pacific Deep Water (NPDW) at 674 m depths near Hawaii using the same extraction method[13]. The NPDW is one of the oldest water masses on earth[14] and thus considered to carry primarily refractory DOM[15]. The reference sample was measured repeatedly (n=12) to assess possible instrument variability and used for comparison with the spectra of the incubated DOM.

Agreement of molecular formulae in mesocosm and NPDW DOM. The ratio of the molecular masses detected in the mesocosm DOM samples to the molecular masses that were also detected in the NPDW DOM reference was calculated on presence-absence basis (Fig. S2.1o). The proportion of shared molecular formulae increased rapidly during the first days and then remained between 45-50%. In the mass range of 280 to 320 Da, however, where the signal intensities of the mesocosm DOM were highest (Fig. S1.1a), as much as 97% of the deep-sea molecular formulae were found in routine analysis mode. For a more detailed look, a mixed sample of mesocosm DOM (M3, days 370, 469, 785) was measured in triplicate at a concentration of 100 mg C L^{-1} and 700 scans were accumulated in a mass window from 380 to 420 Da to increase the sensitivity. The repeated the FT ICR MS analysis in time-consuming, high-sensitivity mode for the mass window (where the NPDW samples had the intensity maximum, Fig. S1.1a), could resolve an 82% agreement of molecular formulae between deep-sea and produced DOM. This finding indicates

that a part of the sample similarity is obscured by the detection limit of the FT ICR MS, inherent to any analytical method.

S1.6 DOM fragmentation experiments

In order to explore structural agreement between assigned molecular formulae detected in the NPDW reference sample and the mesocosm, we fragmented three nominal masses of two homologous series, respectively, in the NPDW reference and a pooled mesocosm sample (M3, days 370, 469, 785). The nominal masses were isolated in the FT ICR MS and fragmented via controlled collision with argon. The collision energy was optimized for each mass in negative mode using ESI ionization, and 150 broadband scans were accumulated per run. The samples were injected at 100 mg C L^{-1}. Fragmentations were conducted in triplicate for each nominal mass. Twelve main fragments (H_2O, CH_4O, $2H_2O$, $2H_2O$, CO_2, CO_2+H_2O, $2CO_2$, $2CO_2$+H_2O, $2CO_2$+CH_4O, $3CO_2$, $2CO_2$+H_2O,$3CO_2$+CH_4O) and their occurrence was compared between NPDW and mesocosm DOM fragmentation of the nominal masses 269, 283, 297, 327, 341 and 365. Peaks with assigned molecular formulae that were detected in all three replicate measurements were considered for data evaluation. Of the 268 identified fragments, 152 (57%) were found in both samples, all 12 main fragments were found in both sample types. Of the 115 fragments that were only detected in one of the samples, 84 were present only in the mesocosm DOM and 31 were present only in the NPDW DOM. Of the 84 exclusive mesocosm DOM fragments, 49 possible fragments were below the detection limit, while 34 fragments were definitely not present in the NPDW sample. All of the exclusive NPDW DOM fragments should have been above detection limit in the mesocosm DOM due to the calculated intensity relative to the mother ion, but were not detected after fragmentation.

S1.7 Estimate of RDOM concentration

Masses with assigned molecular formula that were detected in all FT ICR MS analyses of NPDW DOM were taken into account for the estimate of RDOM concentration in the freshly produced DOM. Each sample was normalized to its sum of peak intensities. Each peak intensity of a sample was then divided by the

respective NPDW DOM peak intensity from which the standard deviation of all NPDW DOM analyses multiplied by $\sigma = 2.576$ had been subtracted to account for some variability, e.g. measurement uncertainty. The ratio was then log-transformed to obtain normal distribution. The number of outliers was determined by counting the log-transformed ratios that deviated from the normal distribution by more than 2.576 times the standard deviation per sample (99% confidence level). This confidence level was chosen because it allowed us to fit ~100% of each NPDW DOM reference analysis into the average NPDW DOM reference. By means of this method, on average 33 masses per sample were excluded from the fit. Next, the maximum percentage of NPDW DOM that could be fitted into the mesocosm DOM was calculated based on normalized peak intensities (Fig. S1.1). Finally, this percentage was multiplied by the solid phase extractable DOC concentration (μmol C L^{-1}) and a factor of 3/2 to account for the solid phase extractability of DOM (~2/3 for RDOM[13]) in order to derive the respective RDOC concentration contained within the mesocosm DOC. This concentration was then used to calculate the percentage of the net community production of DOC (see section S1.8). A proportional behavior of peak intensities and the concentration at which the sample was analyzed could be empirically shown by injecting the NPDW DOM at different concentrations from 4 to 20 mg C L^{-1} amended with the polyethylene glycol internal standard as described in section 1.5 (Fig. S1.2).

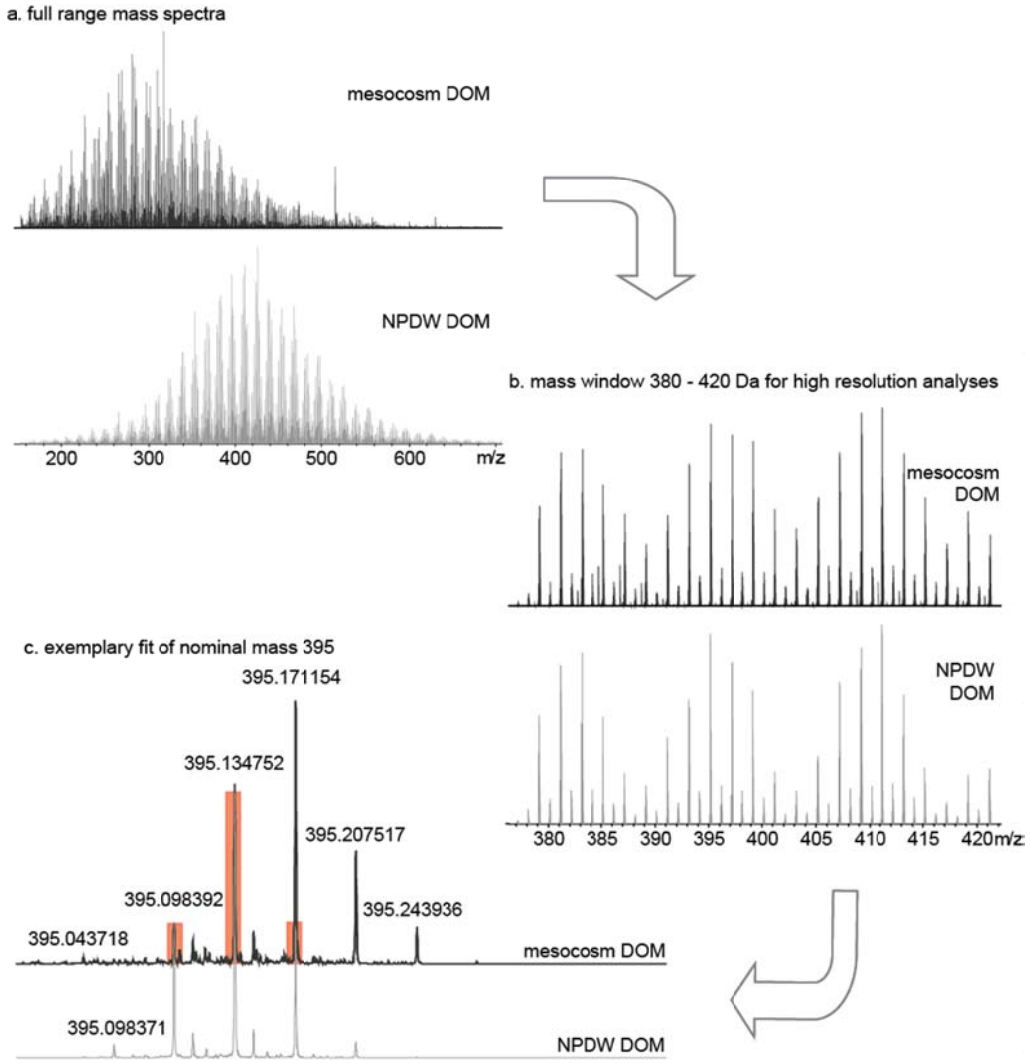

Figure S1.1: Schematic fit of NPDW DOM in mesocosm DOM spectrum. (a) Full range mass spectra of mesocosm DOM (mixture of extracts of M3 from sampling days 370,469, and 785, black) and NPDW DOM (grey) from 150 to 750 Da each illustrate the shift of the intensity maximum towards lower masses of the mesocosm DOM. Routine analysis with acquisition of 500 scans per sample at 20 mg C L^{-1}. (b) High-resolution mass window (380-420 Da) of mesocosm DOM and NPDW DOM measured at increased concentrations (100 mg C L^{-1}) and 700 scans were accumulated per sample. (c) Exemplary fit of NPDW DOM spectrum into mesocosm DOM spectrum at the nominal mass m/z = 395. Highlighted in red is the proportion of the NPDW DOM that can be fitted into the mesocosm DOM spectrum.

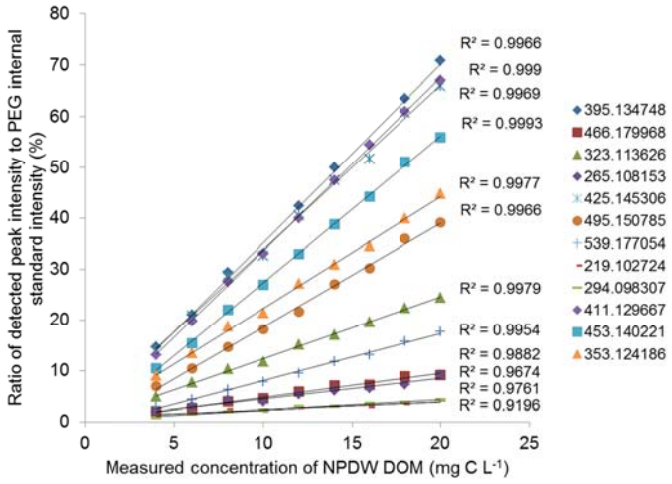

Figure 1.2: Linear correlation of peak intensities and measured concentration in FT ICR MS analyses. The NPDW DOM reference, amended with a polyethylene glycol internal standard, was measured at different concentrations of 4 to 20 mg C L^{-1}. The relative intensity of 12 resolved masses covering the whole mass range and of different relative intensities showed strong linear correlation to the concentration at which the DOM was measured (R^2 values given in plot).

S1.8 Calculation of the net community production (NCP) and comparison with previous studies

The net community production (NCP) constitutes the difference between gross primary production and total respiration of a community. We assumed a ratio of carbon to nitrogen uptake at Redfield stoichiometry[16] (N:C:P = 106:16:1[17]) and calculated the net community production from the uptake of inorganic nitrogen (nitrate and ammonium) or phosphate during the phase of new production until day 56 (final) in our mesocosms:

$$(1)\ NCP = ((NO_3^- initial + NH_4^+ initial) - (NO_3^- final + NH_4^+ final)) * 6.6$$

$$(2)\ NCP = (PO_4^{3-} initial) - (PO_4^{3-} final) * 106$$

The nitrogen-based new production results on average (M1-M3) in a production of 286±5 µmol C L^{-1} through the uptake of 43.1±0.8 µmol L^{-1} inorganic nitrogen in form of nitrate and ammonium. Since nitrogen fixation cannot be ruled out in our

mesocosms, we additionally calculated the phosphorus-based new production: 5.6±0.9 µmol P L^{-1} were assimilated, resulting in a new production of 589±94 µmol C L^{-1}. It has to be noted that the biological drawdown of inorganic carbon and nitrogen was underestimated in some oceanic regions[18] when a Redfield stoichiometry was assumed, but we regard the accuracy as sufficient for the purposes of this study. Among others, one important motivation for the implementation of this study was the fact that previous studies have reported RDOM production rates that exceeded the amounts needed to sustain the global RDOM pool by far (Table 1.3). According to the results presented in this study, a maximum of 0.2 to 0.4% of the net community production of the mesocosms were channeled into the RDOM pool. This result is well within an order of magnitude as the estimates of global RDOM production from the NCP of 0.6%[16].

Table 1.3: Previous estimates of RDOM production. The results of five studies that investigated biological RDOM production in laboratory experiments are summarized. The original amount and type of substrate used as well as the estimated percentage of RDOM production are reported.

Study	Original substrate	Estimated RDOM production as % of original substrate
Brophy & Carlson 1989[19]	equivalents of 415 ng C of leucine and glucose	83 to 99% of produced higher molecular weight material persisted for weeks to months.
Ogawa et al. 2001[20]	208 µmol C L^{-1} as glucose, 132 µmol C L^{-1} as glutamate	5% of glucose- and 7% of glutamate- derived DOC persisted, but also 37 and 50% of the bacterially derived DOC remained and remained largely uncharacterized.
Tranvik 1993[21]	^{14}C labeled glucose (6 µg of DOC L^{-1})	9-15% of originally added total organic carbon remained for 70 days.
Gruber et al. 2006[22]	glucose (10 mmol C L^{-1})	3-5% of initial glucose-C persisted for 36 days of incubation.
Koch et al. 2014	algal exudates (144 µmol C L^{-1}) and glucose (326 µmol C L^{-1})	A calculated minimum 6 µmol C L^{-1} of "non-labile" DOC was produced for algal exudates (~4%), but 115 µmol C L^{-1} (80%) remained after 695 days of incubation. Non-labile DOC production from glucose was estimated at a minimum ~17 µmol C L^{-1} (5%), but 34 µmol C L^{-1} remained after 695 days (~10%).

S2 Supplementary Results

Table S2.1: Abundance of phytoplankton classes over the duration of the experiment. Phytoplankton cell counts for the inoculum (NSW) and the mesocosm M1-M3 until day 469 summarized at class level (10^4 cells L^{-1}).

Sample	**NSW**	**M1**	**M2**	**M3**	**M1**	**M2**	**M3**	**M1**
Day	0	4	4	4	21	21	21	38
Bacillariophyceae	6.6	15.2	4.3	6.0	74.4	1.4	44.9	73.6
Coscinodiscophyceae	152.6	1012.1	770.6	837.8	192.7	222.1	368.6	31.5
Fragilariophyceae	15.4	69.1	11.2	70.6	9.4	0.4	1.3	0.1
Mediophyceae	1.0	0	0.2	0.3	0	0	0	0
Chlorophyceae	0	0	0	0	12.8	71.8	102.5	625.3
Chrysophyceae	0	0	0	0	0	0	0	5.1
Cyanophyceae	0	2665.2	38184	4612.9	4331.0	0	1614.5	2347.5
Cryptophyceae	13.0	9.4	0	2.7	2.7	10.0	10.3	0
Dinophyceae	0.2	0	0	0	0	0	0	0
Euglenophyceae	146.0	1010.9	762.5	834.8	191.4	221.5	368.3	31.5
Protozoa inc. sed.	0	0	0	0	0	15.4	0	0
Prymnesiophyceae	0	217.8	95.6	225.5	53.8	153.8	153.8	4899.9
other Flagellates	69.1	171.7	0	225.5	1829.8	5986.5	5750.8	0

Sample	**M2**	**M3**	**M1**	**M2**	**M3**	**M1**	**M2**	**M3**
Day	38	38	231	231	231	469	469	469
Bacillariophyceae	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1
Coscinodiscophyceae	1309.1	209.0	1.2	392.4	74.1	0	313.9	6.7
Fragilariophyceae	64.7	64.2	0	0	0	0	0	0
Mediophyceae	0	0	0	0	0	0	0	0
Chlorophyceae	0	0	0	0	0	0	0	0
Chrysophyceae	0	0	0	0	0	0	0	0
Cyanophyceae	0	0	0	0	0	0	0	0
Cryptophyceae	0	0	50998	671435	1260.7	563.8	9277.1	1640.1
Dinophyceae	5.1	0	58327	32290	130044	21783	102868	102.5
Euglenophyceae	0	0	0	0	0	0	0	0
Protozoa inc. sed.	64.7	64.2	0	0	0	0	0	0
Prymnesiophyceae/ Flagellates	0	0	0	0	0	0	0	0
other Flagellates	4607.8	515.1	17809	48129	414.8	5459.4	14853	343.4

Table S2.2: Sequencing results of excised bands of DGGE fingerprints. Phylogenetic affiliation, accession number (Acc. no.) of the 16S rRNA phylotypes retrieved in this study from cut DGGE band, and their closest relative, similarity and origin.

Acc. no.	Closest relative (accession no.)	% similarity	Origin
Alphaproteobacteria			
KJ914554	*Rhodobacteraceae bacterium* (KF023497.1)	99	North Sea coastal diatom bloom
KJ914555	*Marivita sp. SSW136* (KC534331.2)	96	shallow water hydrothermal vent, Azores
KJ914556	*Marivita cryptomonadis* (NR_044514.1)	99	Marine phytoplankton *Cryptomonas sp*.
KJ914557	*Marivita cryptomonadis* (NR_044514.1)	99	Marine phytoplankton *Cryptomonas sp*.
KJ914558	*Sulfitobacter sp. KLE1217* (GU644361.1)	99	Marine sediment biofilm, North Atlantic Coast
KJ914559	*Roseovarius sp. M-M10* (JQ739459.1)	99	Seashore sediment, South Korea
KJ914560	*Maribius salinus CL-SP27* (NR_043272.1)	97	Solar saltern in Korea
KJ914561	*Roseovarius crassostreae CV919-312* (AF114484.2)	99	Crassostrea virginica, Maine, USA
Bacteroidetes			
KJ914562	*Uncultured Bacteroidetes bacterium* (HM368331.1)	99	Humic acid-degrading estuarine/marine bacterial communities
KJ914563	*Bacteroidetes bacterium MOLA 411* (AM990678.1)	99	Guanabara Bay, Brazil
Unclassified Sequences			
KJ914564	*Unidentified bacterium K2-S-5 16S* (AY344413.1)	96	Hawaiian Archipelago
KJ914565	*Uncultured bacterium clone REP6-45* (KC853163.1)	97	Surface sea water around decaying *Enteromorpha prolifera*
KJ914566	*Uncultured organism clone SBZO_1899* (JN530474.1)	96	Guerrero Negro hypersaline microbial mat
KJ914567	*Uncultured bacterium clone CBM01B05* (EF395620.1)	99	Seasonally anoxic estuarine waters
KJ914568	*Uncultured bacterium clone D10 NEREIS T270d* (JF774440.1)	98	Bioturbated sediments
KJ914569	*Uncultured bacterium clone G81-44* (HQ601863.1)	97	Biofilm, Great Barrier Reef
KJ914570	*Bacterium HTCC4037* (EF628479)	99	Oregon coast or Bermuda Atlantic Time Series site

Table S2.3: Concentrations of total dissolved carbohydrates (DCCHO) in the mesocosms M1-M3 and M1-M3*dark* on day 1011. DCCHO concentrations are based on the analysis after hydrolysis of the seven neutral monosaccharaides listed.

	M1	M2	M3	M1*dark*	M2*dark*	M3*dark*
	DCCO (µmol L^{-1})					
Fucose	0.2	1.7	0.1	0.0	0.2	0.0
Rhamnose	0.8	1.4	1.1	0.1	0.2	0.4
Arabinose	0.1	0.4	0.1	0.0	0.0	0.0
Galactose	0.4	11.4	0.1	0.1	0.1	0.0
Glucose	0.9	6.7	0.9	0.2	0.1	0.2
Mannose	0.1	2.5	0.2	0.0	0.0	0.0
Xylose	0.3	1.8	0.1	0.1	0.2	0.0
Sum	2.8	25.9	2.7	0.6	0.8	0.7

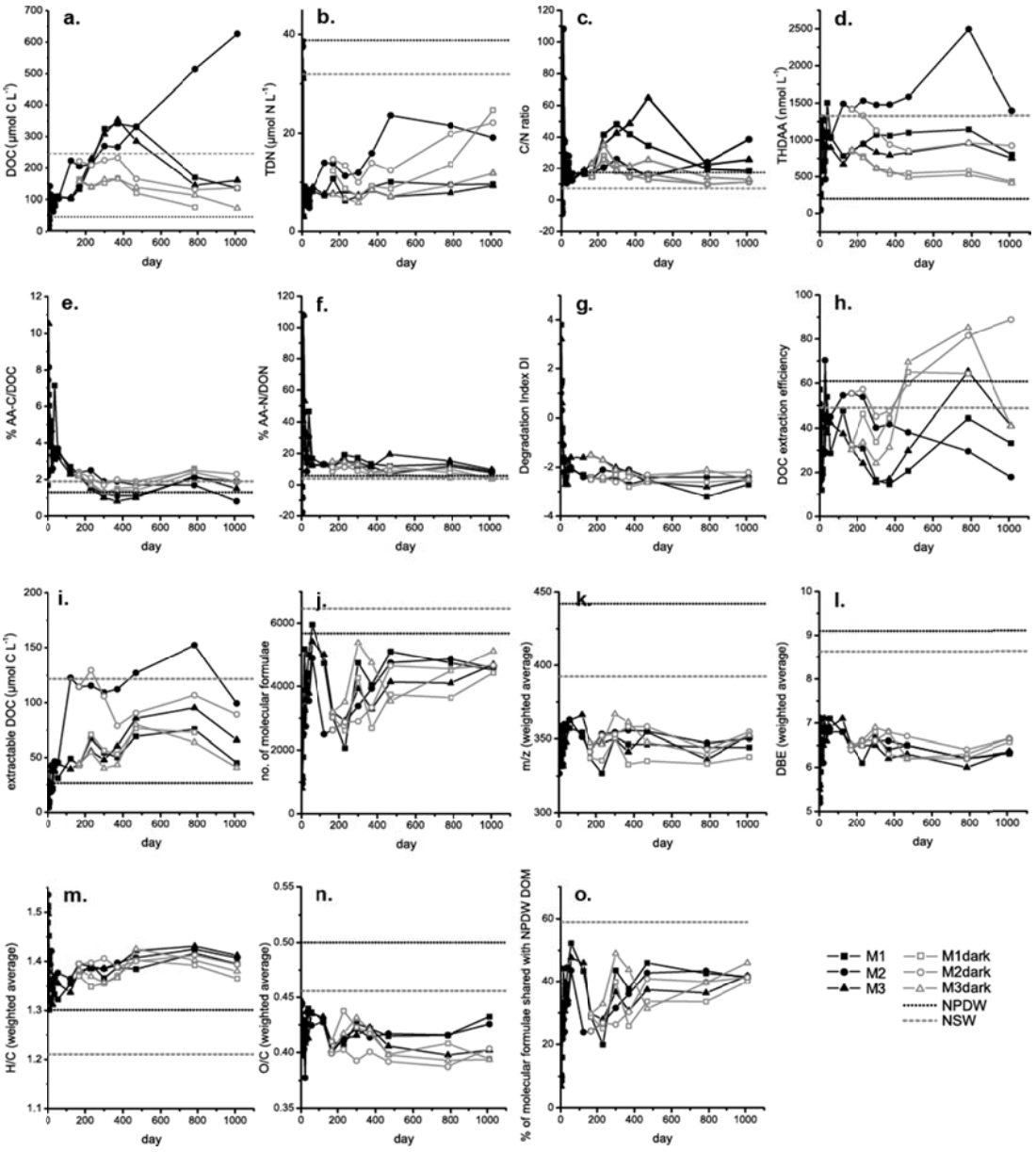

Figure S2.4: Bulk and molecular parameters over the duration of the experiment. (a) DOC concentration, (b) TDN concentration, (c) C/N ratio calculated from DOC/DON, (d) THDAA concentration, (e) ratio of THDAA-derived carbon to DOC, (f) ratio of THDAA-derived nitrogen to DON, (g) the degradation state index[9] calculated from relative contributions of certain amino acids, (h) extraction efficiency on carbon basis, (i) concentration of extractable carbon. Molecular DOM composition derived from FT ICR MS analyses: (j) number of assigned molecular formulae, (k) peak intensity-weighted molecular mass, (l) peak intensity-weighted double bond equivalent DBE, (m) peak intensity-weighted hydrogen-to-carbon ratio H/C, (n) peak intensity-weighted oxygen-to-carbon ratio O/C based on assigned molecular formulae (o) percentage of molecular formulae shared with NPDW DOM, calculated from the total number of assigned formulae per sample.

Fig. S2.4cont. North Sea Water reference (NSW) values determined in the inoculum of the mesocosms, North Pacific Deep Water reference (NPDW) values are taken from Green et al. 2014[13] or literature compiled in Hansell & Carlson 2002[23] for the deep ocean. M1-M3 denote light mesocosms, M1-M3*dark* denote dark mesocosms.

S3 Supplementary References

1 Zobell, C. E. Studies on marine bacteria. I. The cultural requirements of heterotrophic aerobes. *Journal of Marine Research* **4**, 42-75 (1941).

2 Gasol, J. M. & Del Giorgio, P. A. Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. *Scientia Marina* **64**, 197-224 (2000).

3 Giebel, H. A., Brinkhoff, T., Zwisler, W., Selje, N. & Simon, M. Distribution of Roseobacter RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. *Environ Microbiol* **11**, 2164-2178 (2009).

4 Muyzer, G., Dewaal, E. C. & Uitterlinden, A. G. Profiling of Complex Microbial-Populations by Denaturing Gradient Gel-Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes-Coding for 16s Ribosomal-Rna. *Appl. Environ. Microbiol.* **59**, 695-700 (1993).

5 Osterholz, H., Dittmar, T. & Niggemann, J. Molecular evidence for rapid dissolved organic matter turnover in Arctic fjords. *Mar. Chem.* **160**, 1-10 (2014).

6 Laskov, C., Herzog, C., Lewandowski, J. & Hupfer, M. Miniaturized photometrical methods for the rapid analysis of phosphate, ammonium, ferrous iron, and sulfate in pore water of freshwater sediments. *Limnology and Oceanography-Methods* **5**, 63-71 (2007).

7 Schnetger, B. & Lehners, C. Determination of nitrate plus nitrite in small volume marine water samples using vanadium(III)chloride as a reduction agent. *Mar. Chem* (2014).

8 Lunau, M., Lemke, A., Dellwig, O. & Simon, M. Physical and biogeochemical controls of microaggregate dynamics in a tidally affected coastal ecosystem. *Limnology and Oceanography* **51**, 847-859 (2006).

9 Dauwe, B. & Middelburg, J. J. Amino acids and hexosamines as indicators of organic matter degradation state in North Sea sediments. *Limnology and Oceanography* **43**, 782-798 (1998).

10 Mopper, K. *et al.* Determination of sugars in unconcentrated seawater and other natural waters by liquid chromatography and pulsed amperometric detection. *Environmental Science & Technology* **26**, 133-138 (1992).

11 Dittmar, T. The molecular level determination of black carbon in marine dissolved organic matter. *Organic Geochemistry* **39**, 396-407 (2008).

12 Koch, B. P., Dittmar, T., Witt, M. & Kattner, G. Fundamentals of Molecular Formula Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter. *Analytical Chemistry* **79**, 1758-1763 (2007).

13 Green, N. W. *et al.* An intercomparison of three methods for the large-scale isolation of oceanic dissolved organic matter. *Mar. Chem*, (2014).

14 Stuiver, M., Quay, P. D. & Ostlund, H. G. Abyssal Water Carbon-14 Distribution and the Age of the World Oceans. *Science* **219**, 849-851 (1983).

15 Hansell, D. A. Recalcitrant Dissolved Organic Carbon Fractions. *Annual Review of Marine Science* **5**, 421-445 (2013).

16 Hansell, D. A. & Carlson, C. A. Net community production of dissolved organic carbon. *Global Biogeochemical Cycles* **12**, 443-453 (1998).

17 Redfield, A. C. in *James Johnstone Memorial Volume* (ed R.J. Daniel) 177-192 (University Press, 1934).

18 Sambrotto, R. N. *et al.* Elevated consumption of carbon relative to nitrogen in the surface ocean. *Nature* **363**, 248-250 (1993).

19 Brophy, J. E. & Carlson, D. J. Production of biologically refractory dissolved organic carbon by natural seawater microbial populations. *Deep Sea Research Part A. Oceanographic Research Papers* **36**, 497-507 (1989).

20 Ogawa, H., Amagai, Y., Koike, I., Kaiser, K. & Benner, R. Production of Refractory Dissolved Organic Matter by Bacteria. *Science* **292**, 917-920 (2001).

21 Tranvik, L. J. Microbial transformation of labile dissolved organic matter into humic-like matter in seawater. *Fems Microbiology Ecology* **12**, 177-183 (1993).

22 Gruber, D. F., Simjouw, J.-P., Seitzinger, S. P. & Taghon, G. L. Dynamics and Characterization of Refractory Dissolved Organic Matter Produced by a Pure Bacterial Culture in an Experimental Predator-Prey System. *Appl. Environ. Microbiol.* **72**, 4184-4191 (2006).

23 Hansell, D. A. & Carlson, C. A. *Biogeochemistry of marine dissolved organic matter*. (Academic Pr, 2002).

4 Terrigenous input and microbial processing - driving forces of dissolved organic matter composition in the North Sea

Helena Osterholz[1], Gabriel Singer[2], Jutta Niggemann[1],
Meinhard Simon[3], Thorsten Dittmar[1]

[1]Research Group for Marine Geochemistry (ICBM-MPI Bridging Group), Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky University Oldenburg, Carl-von-Ossietzky-Strasse 9-11, 26129 Oldenburg, Germany

[2]Leibniz-Institute of Freshwater and Inland Fisheries (IGB), Müggelseedamm 310, 12587 Berlin, Germany

[3]Biology of Geological Processes, Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky University Oldenburg, Carl-von-Ossietzky-Strasse 9-11, 26129 Oldenburg, Germany

4.1 Abstract

The North Sea as marginal sea of the North Atlantic is largely surrounded by land and receives most of the riverine discharge of the central European drainage system. High loads of dissolved organic matter (DOM) carried by the rivers together with autochthonous production and degradation processes heavily impact the carbon budget of the confined ocean. The main drivers of bulk dissolved organic carbon (DOC) concentration and molecular DOM as well as microbial community composition were identified in the water column along a transect from the island Helgoland towards the Norwegian coast and in the Skagerrak area. DOC concentrations ranged between 60 and 170 µmol C L^{-1}, reflecting the conservative mixing of high DOC freshwater with Atlantic Water masses from the North. The DOM quality, analyzed via ultrahigh resolution mass spectrometry, was heavily impacted by terrigenous DOM compounds such as dissolved black carbon and polyphenolic compounds, which decreased from ~27% to ~17% of all detected molecular formulae along the salinity gradient. Besides salinity, strong correlations were revealed between the total microbial community (DNA-based) and molecular DOM composition, whereas the active microbial community (RNA-based) only exhibited a weak link to the DOM molecular composition. We hypothesize that while the modifications of the DOM through the active microbial community rather affect the labile DOM pool that is rapidly turned over and cannot completely be captured by the applied methods, the semi-labile DOM and total microbial community retain the history of the water masses on similar timescales of several months.

4.2. Introduction

The North Sea is a 750,000 km^2 large marginal sea of the Atlantic Ocean bordered by Great Britain, Norway, Denmark, Germany, the Netherlands, Belgium and France. The main water exchange with the Atlantic Ocean (51,000 km^3 yr^{-1} inflow and 56,700 km^3 yr^{-1} outflow) occurs via the northern boundary. The water exchange through the Dover Strait between the Netherlands and Great Britain contributes

~4,900 km^3 yr^{-1} (Eisma and Kalf (1987). Compared to these numbers, the freshwater flow into the system consisting of precipitation, freshwater input from rivers, and the outflow from the Baltic Sea in the East is relatively small and sums up to 895-955 km^3 yr^{-1} (Radach and Pätsch 2007; Thomas et al. 2005). The major discharging rivers are the Rhine and the Elbe, which together transport water from a large part of the central European drainage basin. These two rivers are responsible for a significant input of pollutants, nutrients and organic material into the North Sea (Radach and Pätsch 2007). Dissolved organic matter (DOM), the dominant form of organic matter transported by the rivers (Wetzel 1984), is an important source of energy and nutrients (N, P) for coastal microbial metabolism and primary production (Azam et al. 1983; Tranvik 1992). The input of nutrients is traditionally underestimated as organic forms of nutrients are infrequently quantified, but may constitute the prevailing form of dissolved terrestrial nutrient supply to the ocean (Guo et al. 2004).

The huge amounts of DOM carried by rivers are either derived from autochthonous primary production or of terrigenous origin. From a molecular perspective, terrigenous components are mainly lignin phenols derived from constituent compounds of vascular plants (Hernes and Benner 2006; Opsahl and Benner 1997) and pyrogenic black carbon, a product of biomass burning (Dittmar and Paeng 2009; Jaffé et al. 2013) with rather distinct composition and structural properties. Mobilized mainly from soils and subsequently transported to the oceans by rivers, terrigenous DOM contributes between <1% of the dissolved organic carbon (DOC) in the Pacific to 33% of the DOC in the Arctic Ocean (Hernes and Benner 2002; Opsahl and Benner 1997; Opsahl et al. 1999). The fate of terrigenous DOM in the ocean is still unclear. Especially in coastal areas, which are the first receivers of terrigenous DOM, remineralization is a possible source of nutrients for primary production (Moran and Hodson 1994; Opsahl and Benner 1997). Further, in the shallow coastal waters, photodegradation reduces the aromaticity of condensed aromatic compounds (Opsahl and Benner 1997; Stubbins et al. 2010), which has been proposed to render DOM more susceptible to microbial consumption (Benner and Ziegler 2000; Miller and Moran 1997). Also, redox interfaces at oxic-anoxic

transitions in the sediments of e.g. estuaries or tidal flats may play an important role in molecularly modifying the flux of DOM from land to ocean by selectively trapping vascular plant-derived and pyrogenic DOM (Riedel et al. 2013).

In addition to the allochthonous input of DOM to the North Sea through rivers, groundwater discharge or aeolian deposition, and primary production in the marine environment may add substantial amounts of autochthonous DOM (Hedges 1992), thereby affecting concentration and composition of the marine DOM pool. Besides autotrophic production, the interactions between heterotroph organisms and DOM are numerous and further alter DOM composition. Conceptually based in the microbial loop (Azam 1998) and partly in the microbial carbon pump (Jiao et al. 2010), these processes include selective consumption for growth, energy storage and metabolism, transformation, production of secondary compounds and cleavage products as well as excretion. In order to understand these interactions, it is important to identify the key players on both sides, the composition of the DOM and the microbial community (hereafter MC) structure, in detail.

For the North Sea and the German Bight and based on samples collected during the same cruise as this study, Wemheuer et al. (in prep.) showed that bacteria of the orders *Flavobacteriales* and *Rhodobacterales* were most abundant in surface waters during the summer. They assessed MC composition using 454 pyrosequencing of RNA and DNA with the aim to differentiate actively growing taxa from a total community potentially including numerous inactive taxa. In this study, *Alphaproteobacteria*, especially the different clusters of the *Roseobacter* clade, were found to be most abundant and active, followed by *Bacteroidetes* and *Cyanobacteria*. *Alphaproteobacteria* have already previously been found to dominate the MC in the North Sea during spring, i.e., before the onset of the annual phytoplankton bloom (Teeling et al. 2012). During the bloom, the number of *Bacteroidetes* strongly increased, while *Gammaproteobacteria* were dominating during later stages of algal decay. The heterotrophic *Bacteroidetes* are known to be important degraders of high molecular weight organic matter in the ocean and preferably grow attached to particles (e.g. Fernandez-Gomez et al. 2013). In

contrast, *Proteobacteria*, especially the most abundant and highly diverse *Roseobacter* clade of the *Alphaproteobacteria* subclass, preferentially consume monomers and live freely suspended in the pelagial (Buchan et al. 2005; Giebel et al. 2011). There is growing evidence that the MC composition in the marine realm is driven to a large extent by deterministic selection (Stegen et al. 2012; Valentin-Vargas et al. 2012) through environmental factors, among which substrate availability and composition of DOM rank prominently. Closely related and functionally similar taxa are found in similar habitats more often than expected by chance (Kraft et al. 2007) and substrate-specific allocation to the vast niche space offered by complex DOM may contribute to maintaining the highly diverse microbial communities in the oceans (Carlson et al. 2007; Zinger et al. 2011) as it has been proposed for phytoplankton (Hutchinson 1961).

Technical advances in molecular microbial ecology (Edwards and Dinsdale 2007) and organic geochemistry (Koch et al. 2005; Kujawinski 2002) have made it possible to analyze both, the microbial community and the dissolved organic matter, at similarly high resolution. Pyrosequencing facilitates the classification of 100,000s of DNA sequences below phylum level and FT ICR MS (Fourier Transform Ion Cyclotron Resonance Mass Spectrometry) typically yields tens of thousands of molecular formulae for an oceanic DOM sample. Linking this detailed information from molecular microbiology and organic geochemistry is a promising step towards the understanding of the role of microorganisms in the global marine carbon cycle. A major challenge is the combination and interpretation of the complex information obtained by the ultrahigh resolution methods via multivariate statistics.

In this study we aim to (i) characterize the molecular composition of the North Sea DOM pool along a gradient of terrestrial influence starting from Helgoland up to ~60°N, and (ii) identify environmental variables as well as features of the MC associated with the gradient of DOM composition. We hypothesize that the MC composition, especially the active community, is closely related to the composition of their sustaining energy source (Azam et al. 1983), the DOM.

4.3 Materials & Methods

Study area and sampling. Water samples were obtained along a transect of 15 stations from the German coast towards Norway on R/V Heincke cruise HE361 from July 7 to July 29 2011 (Fig. 1). Additionally, 3 stations (stations 8, 10 and 27) were sampled in the Skagerrak area. At all stations, surface water was collected at 3-4 m depth using 5 L Niskin bottles attached to a CTD; additional sampling depths were selected depending on salinity, water temperature and fluorescence data reported by the CTD. In total, our study encompasses 45 samples (see supplementary material). Chlorophyll *a* and phaeopigment concentration, glucose and dissolved free amino acid turnover rates as well as total bacterioplankton cell counts were aquired (Wemheuer et al. in prep.). The microbial community (MC) composition was analyzed for the surface water of 12 stations by analysis of 16S rRNA amplicons of DNA and RNA samples as described in Wemheuer et al. (in prep.).

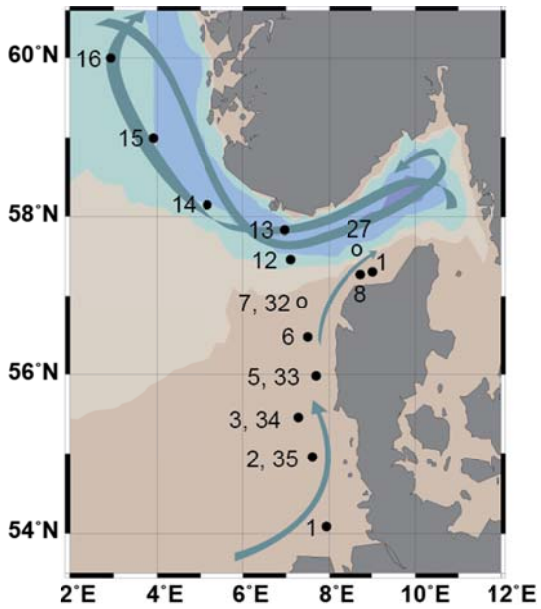

Fig. 1: Sampling stations in the German Bight and North Sea. Circles denote sampling stations, filled circles indicate stations where RNA/DNA-based pyrosequencing data is available for surface samples, and arrows indicate approximate directions of the main currents after Turrell (1992). Map created with Ocean Data View (R. Schlitzer, http://odv.awi.de).

DOC/TDN quantification and solid phase extraction. Water samples for DOM analysis were filtered through GF/F filters (precombusted 400°C, 4 hrs, Whatmann, United Kingdom) and acidified to pH 2 with 25 % HCl (p.a. grade, Carl Roth, Germany). Triplicate subsamples for dissolved organic carbon (DOC) and total dissolved nitrogen (TDN) analysis were stored at 7°C in the dark pending laboratory analyses. DOC and TDN concentrations were analyzed by high temperature catalytic combustion using a Shimadzu TOC-VCPH/CPN Total Organic Carbon Analyzer equipped with an ASI-V autosampler and a TNM-1 module. TDN reported in this study encompasses both organic and inorganic fractions. Prior to analysis, the acidified samples were purged with synthetic air to remove dissolved inorganic carbon. L-arginine solutions ranging from 5 to 500 µmol C L^{-1} and 6.6 to 333.3 µmol N L^{-1}, respectively, were used for calibration and Deep Atlantic Seawater reference material (DSR, D.A. Hansell, University of Miami, Florida, USA) was measured to control for instrumental precision and accuracy. The standard deviation of triplicate sample analyses was on average 2 µmol C L^{-1} for DOC and 1 µmol N L^{-1} for TDN analysis. According to the DSR analyses (n=10), the precision was 2.8±0.9% for DOC and 3.3±1.1% for TDN, while the accuracy was 5.1±4.2% for DOC and -4.8±1.2% for TDN, respectively. For mass spectrometry, DOM was extracted from 2 liters of filtered and acidified seawater with commercially available modified styrene divinyl benzene polymer cartridges (PPL, Varian, USA) as described in Dittmar et al. (2008). After loading the sample onto the adsorber, cartridges were rinsed with acidified ultrapure water (pH 2, HCl 25%, p.a., Carl Roth, Germany) to remove salts, dried by flushing with nitrogen gas, and eluted with 6 ml of methanol (HPLC-grade, Sigma-Aldrich, USA). Extracts were stored in amber vials at -20°C. The extraction efficiency was on average 52±5% on carbon basis.

Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT ICR MS). The mass spectra were obtained on a 15 Tesla Solarix FT ICR MS (Bruker Daltonics, USA) equipped with an electrospray ionization source (Bruker Apollo II) applied in negative mode. DOM extracts were diluted to a final DOC concentration of 20 mg C L^{-1} in a 1:1 mixture of ultrapure water and methanol (HPLC-grade,

Sigma-Aldrich, USA). A total of 500 scans were accumulated for each sample, the mass window was set to 150-2000 Da. The spectra were calibrated with an internal calibration list containing more than 50 known peaks with identified molecular formulae spanning the mass range between 281 and 621 Da using the Bruker Daltonics Data Analysis software package. The mass to charge ratio, resolution and intensity were then exported and processed using in-house Matlab routines. Detected peaks were matched over the whole dataset and molecular formulae (maximum elemental abundances set to $C_nH_nO_nN_4S_2$) were assigned to peaks with a minimum signal-to-noise ratio of 4 and based on a maximum error of 0.5 ppm on the mass axis. CH_2 homologous series were calculated and longer homologous series extending to lower masses were considered correct for multiple assignments. The modified Aromaticity Index (AI_{mod}=(1+C-½O-S-½H)/(C-½O-S-N-P)) was calculated to assess the presence and extent of aromatic structures (Koch and Dittmar 2006). Based on the elemental composition, molecules detected via FT ICR MS were assigned to "molecular categories" (black carbon, polyphenols, highly unsaturated, unsaturated, aliphatics, saturated fatty acids, sugars, peptides) following the approach of Šantl-Temkiv et al. (2013).

Statistical Evaluation. Total and active MC composition obtained by pyrosequencing-based analysis of 16S rRNA amplicons from environmental DNA and RNA were provided by Wemheuer et al. (in prep). In short, the acquired sequences were clustered at 1% similarity (UCLUST algorithm, Edgar 2010) and then grouped to operational taxonomic units (OTUs) according to phylogenetic affiliation. Six resolutions from phylum to approximately species level were each tested in the following analyses. MC composition was expressed in terms of relative abundances (as fractions of 1) of sequences for each of the 6 phylogenetic resolutions and RNA as well as DNA. Similarly, the DOM composition was assessed based on the relative peak intensities of the FT ICR MS spectra, i.e. absolute peak intensities normalized to the total sum of peak intensities of a given spectrum. Only peaks with assigned molecular formula were taken into account (n=6542) in two datasets: DOM45 included all measured samples, DOM12 included the DOM analyses at the 12 stations for which pyrosequencing data was available.

To describe gradients of compositional change in the MC and in DOM, we transformed matrices of relative abundance (or intensity, respectively) into separate dissimilarity matrices based on Bray-Curtis compositional dissimilarity (Bray and Curtis 1957; Legendre and Legendre 1998). For each dissimilarity matrix, principal coordinate analysis (PCoA) was then used to graphically describe relationships among sites (i.e., samples) based on the first two major axes of variation. We opted for PCoA as a means to show major variation the datasets (instead of non-metric multidimensional scaling, NMDS) to make this step of our analyses compliant with the following canonical analyses; PCoA and NMDS (data not shown) did not result in visually obviously differing ordinations of sites. Canonical analyses targeted (i) the relationships of MC or DOM composition to environmental factors and (ii) the linkage between MC and DOM compositions. The first was achieved by running a canonical correlation analysis (CCorA) using any of the PCoA-results and a chosen environmental factor. This analysis aims at identifying a gradient of compositional change (of MC or DOM) which most strongly correlates with the environmental factor, the analysis is also known as a canonical analysis of principal coordinates (CAP, Anderson and Willis 2003) and offers the possibility of a formal hypothesis test by permutation. We here note that CAP of, for instance, DOM with an environmental variable may yield a significant result, while the same environmental variable is not necessarily significantly correlated to first principal coordinates of DOM. This merely indicates that a relationship of DOM composition with the environmental variable must be rather subtle and does not express itself in major compositional changes of DOM; in such a case a canonical analysis (as CAP) offers the necessary power to detect the relationship, which, however, may remain undetected in an unconstrained, descriptive analysis as PCoA, To link MC composition with the composition of DOM we followed an approach similar to CAP, but needed to first select a subset of principal coordinates for both involved dissimilarity matrices (one for MC, one for DOM) to avoid overfitting; this was done based on scree plots and with the aim to capture at least 70% of the total variation of each dataset. We then fed the selected subsets of principal coordinates into a CCorA aiming to identify the first canonical axes pair with maximized

correlation between MC composition and the composition of DOM. Significance of canonical correlation was again computed by permutation. To check if the subset of principal coordinates was able to capture the relevant parts of variation contained in the full dissimilarity matrices we added a Mantel test checking for a significant association of MC composition with the composition of DOM. While the latter offers a global hypothesis test of association, the canonical analysis step is able to quantitatively map this association onto dimensions in the spaces of MC and DOM, i.e., the canonical axes pair. As a last step we (i) correlated relative peak intensities with the canonical axis in DOM space and color-coded sum formulas in Van-Krevelen plots according to this correlation, and (ii) correlated relative abundances with the canonical axis in MC space and color-coded OTU names on a phylogenetic tree according to this correlation. All multivariate statistics were performed in R (version 2.15.2, R Development Core Team 2012, [http://cran.r-project.org/]) and using the package "vegan" (Oksanen et al. 2012).

4.4 Results & Discussion

Sources and distribution of DOC and TDN in the North Sea. Salinity along our 700 km North Sea study transect ranged from 29.8 psu at the southernmost station (station 1) to 35.4 psu at almost 60°N (station 16) with slight vertical stratification (<5 psu difference between surface and bottom water, on average: 1.3±1.6 psu). Exceptions were the tidally impacted areas south of 56°N and the shallow Skagerrak (stations 8, 10, 27). Our results are readily interpretable by North Atlantic water flowing in at the northern boundary and circling through the North Sea (Thomas et al. 2005), and an influence from less saline Baltic Water in the Skagerrak (Danielssen et al. 1996). DOC concentrations ranged between 68 and 170 µmol C L^{-1} and correlated significantly with salinity (Pearson's correlation: $r=-0.94$, $p<0.001$, Fig. 2a). This finding indicates conservative mixing of high-DOC freshwater, largely consisting of riverine input from Rhine, Meuse and Scheldt in the Netherlands, as well as the Elbe, Weser and Ems at the German coast (Radach and Pätsch 2007), and Atlantic water with lower DOC concentrations (Hansell and

Carlson 1998). Fully conservative mixing would yield a DOC concentration of 630 µmol L^{-1} at a salinity of 0 (y=-15.873x+630.51, calculated from measured concentrations, this study) which is well within the range of the Ems (570 to 675 µmol C L^{-1}), but higher than the 340-380 µmol C L^{-1} and the 240-440 µmol C L^{-1} reported for Elbe and Rhine, respectively (Abril et al. 2002). According to Thomas et al. (2005), the Baltic inflow impacts the North Sea water salinity and DOC content in the area along a narrow strip along the Norwegian coast; in our study this is merely reflected in the elevated DOC concentration of 149 µmol L^{-1} at the low salinity (30.06 psu) surface water of station 13 located in the Norwegian Trench. At the oceanic end of our investigated gradient (station 16), we find the lowest measured DOC concentration of 54 µmol C L^{-1}, which is still higher than the concentrations reported for the Greenland Sea Deep Water with 48 µmol C L^{-1} or the North Atlantic Deep Water with 45 µmol C L^{-1} (Hansell and Carlson 1998). This supports the hypothesis by Thomas et al. (2003) who suggested the North Sea to represent a "bypass pump" that increases the carbon content of the Atlantic Ocean water during its circulation. For the temporarily stratified central North Sea previous studies are limited, but Suratman et al. (2009) report DOC concentrations between 68 and 318 µmol L^{-1}, with the highest values occurring during the phytoplankton blooms in spring and summer. In the tidally mixed southern North Sea, the influence of riverine input increases and DOC concentrations range between 83 and 417 µmol L^{-1} (Cadée 1982).

We measured TDN concentrations ranging between 2.5 and 20.5 µmol N L^{-1} (Fig. 2b) with many surface stations being depleted in N, most likely due to removal during recent phytoplankton blooms (Colijn and Cadée 2003). The water depths between 100 and 350 m were enriched in TDN, which can be attributed to the lack of active autotrophs and dominance of remineralization processes at these depths (Herndl and Reinthaler 2013; Shaffer 1996). Mean total N concentrations, including particulate nitrogen, of the rivers discharging into the North Sea were calculated to range between 300 and 400 µmol N L^{-1} (Brion et al. 2004), but TDN concentrations in the Elbe estuary were reported at 72.2±17.6 µmol N L^{-1} (Kerner and Spitzy 2001). Coastal ecosystems in general are considered to be important sinks for nutrients

(Nixon et al. 1996), and in the case of the North Sea they presumably alleviate the flux of N to the North Atlantic via uptake by autotroph and heterotroph microorganisms as well as by denitrification (Brion et al. 2004). During our sampling campaign, however, we detected no extensive phytoplankton blooms. A few stations were found with deep chlorophyll maxima between 18-23 m depth, but chlorophyll *a* concentrations were generally low (<3.3 µg L^{-1}, av. 1.3 µg L^{-1}, see Wemheuer et al. in prep) when compared to the annual cycles shown in Radach and Pätsch (1997), where concentrations above 10 µg L^{-1} are reached periodically.

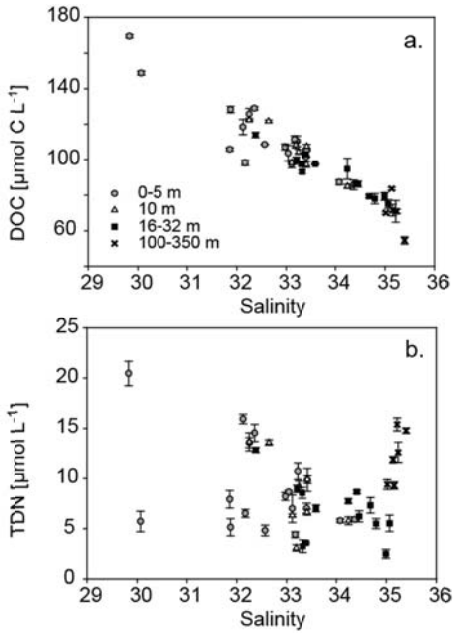

Fig. 2: DOC (a) and TDN (b) concentrations (±standard deviation of triplicate analyses) versus salinity grouped by depth (0-5 m n=18, 5-10 m n=8, 16-32 m n=13, 100-350 m n=6) at all stations.

Molecular DOM composition: tracing terrigenous input. Via FT ICR MS we were able to resolve 10,198 different masses in all 45 DOM extracts. On average molecular formulae were assigned to 68±1% of the masses, excluding ^{13}C isotopic peaks. Surface water at stations 1 near Helgoland and 13 close to Norway with salinities of 29.8 and 30, respectively, represent the freshest water masses on the transect while at the northernmost station 16 the salinity is above 35 at all depths. The average intensity-weighted size of the molecules increased from 384 Da to

420 Da from the lowest to the highest salinity (Pearson's correlation: $r=0.43$, $p<0.05$). Accordingly, we detected changes along the salinity gradient within the molecular classes of DOM (Fig. 3). The percentage of black carbon and polyphenol-type molecules decreased with increasing salinity, the modified aromaticity index similarly decreased with increasing salinity (-0.81, $p<0.05$, data not shown). In contrast, the fraction of highly unsaturated molecules, peptides and sugars decreased with increasing salinity (Fig. 3), while the contributions of saturated fatty acids and unsaturated aliphatics to the total formulae count remained constant and did not show significant gradients with salinity. Black carbon, a product of incomplete combustion (Goldberg 1985; Jaffé et al. 2013), and polyphenolic compounds, which occur only in vascular plants, are unambiguous tracers of terrigenous DOM. Based on the sum of the intensity-weighted percentages of the molecular classes black carbon and polyphenols we estimate the fraction of terrigenous DOM as 10.8% of the total DOC at station 1, the southernmost station close to the island of Helgoland. Even though this estimate is based on only two compound classes, the decrease to 3.6% at the northernmost station with the lowest DOC concentration of only 54 $\mu mol\ L^{-1}$ compares well to the findings of Hernes and Benner (2006). These authors describe a terrigenous contribution to DOC in the North Atlantic of 1-2% as quantified by the analysis of lignin phenols in the water column.

Canonical Analysis of Principal Coordinates (CAP) with all 45 DOM samples and using salinity as a constraint revealed a highly significant ($p<0.001$) association of the molecular composition of DOM with salinity across the North Sea sampling transect (Fig. 4a). This finding corresponds well to the previously described largely conservative mixing of the low-salinity coastal water masses with Atlantic Water in the North Sea and establishes salinity, i.e. the conservative mixing process itself, as the main determinant of DOM composition. However, 47% of the variability remains unexplained, pointing to additional factors that influence molecular DOM composition, e.g. microbial interactions, including production, degradation and transformation processes.

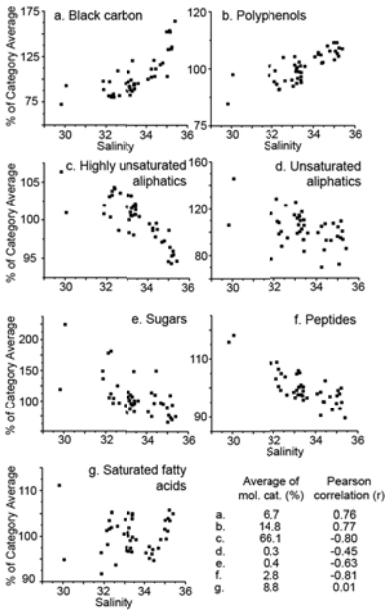

Fig. 3: Variation of molecular formulae contributing to molecular category in % of average along the salinity gradient. The average contribution to the total number of molecular formulae and the Pearson's correlation to salinity are given in the lower right corner.

Since only 12 samples were available to investigate the microbial imprint on the DOM, this analysis was repeated with the reduced dataset (Fig. 4b). None of the measured environmental parameters such as chlorophyll *a* concentration, bacterial production, temperature, etc. could be confirmed as the main factor influencing DOM composition on the first PCoA axis which accounted for 33% of the variability, but salinity still explained 29% of the DOM variability on the second PCoA axis (Fig. 4b). The weaker influence of salinity on the DOM composition most likely results from the fact that the reduced dataset only included surface samples in a narrower salinity range (29.8-33.4 psu).

Linking bacterial and DOM communities – choosing the phylogenetic resolution.

The MC composition of 12 of the 13 stations described in Wemheuer et al. (in prep.) was used for the calculation of coherences between molecules and microbes on basis of the total (DNA-based) and active (RNA-based) bacterial community composition. *Alphaproteobacteria* and *Flavobacteria* were most abundant at all stations. The contribution of *Cyanobacteria* and *Gammaproteobacteria* was patchy, whereas

Cytophagia, *Opitutae* and *Deltaproteobacteria* accounted for only minor fractions of the MC. The overall trends in MC composition were often similar between DNA- and RNA-based analyses, but relative contributions of the subgroups at lower phylogenetic levels to the community structure varied considerably.

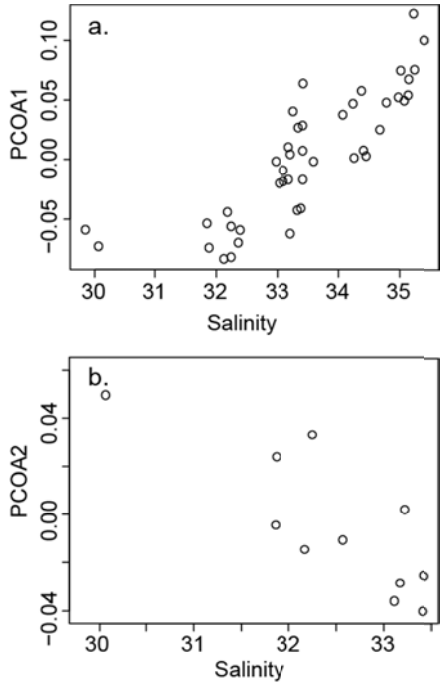

Fig. 4: PCoA1 eigenvalues (a, including 53% of the variability of the dataset) plotted against salinity (r=0.99, p<0.001) for all 45 DOM samples. PCoA2 eigenvalues (b, including 28% of the variability of the dataset) plotted against salinity (r=0.99, p<0.001) for the reduced dataset with 12 DOM samples.

All statistical analyses concerning the MC were then performed for the six levels of phylogenetic resolution approximately equivalent to phylum to species levels (Tab. 1a). We found a higher number of OTUs for the active community when compared to the total MC with a DNA/RNA ratio of 63-85%, denoting that a significant portion of the rare taxa may be proportionally more active (Campbell et al. 2011; Jones and Lennon 2010). The percentage of the OTUs that appeared in both, RNA and DNA datasets, was usually <54%, but always >36%. A Mantel test confirmed a strong correlation between DNA and RNA based community composition ($p<0.01$, data not shown), that has previously been shown by Campbell et al. (2011) for the coastal waters off Virginia, USA. Relative abundances of the RNA and DNA OTUs

are overall not significantly correlated to any environmental variable except salinity at any phylogenetic level according to Constrained Analysis of Principal Coordinates (CAP).

Tab. 1: Overview of (a) OTUs per phylogenetic level for RNA- and DNA-based, percentage of OTUs in DNA-based dataset which were also found in the RNA-based dataset (and vice versa), number of OTUs found in both datasets (percentage computed from total number of OTUs found by either approach); (b) Results of canonical correlation analysis and redundancy per phylogenetic level; (c) Redundancy of DNA on DOM and RNA on DOM, as well as correlation and significance of the correlation of the first canonical axes.

a.

phylogenetic level	# of DNA OTUs	# of RNA OTUs	%DNA/RNA	%RNA/DNA	# of OTUs in both RNA and DNA datasets
phylum	17	22	82	64	14 (54%)
class	40	50	78	62	31 (50%)
order	105	124	74	63	78 (50%)
family	192	240	76	60	145 (48%)
genus	372	513	74	54	277 (44%)
species	599	946	71	45	423 (36%)

b.

phylogenetic level	DNA CCorA	RNA CCorA
phylum	0.60	0.030*
class	0.049*	0.440
order	0.011*	0.067
family	0.014*	0.16
genus	0.0046***	0.046*
species	0.0007***	0.028*

c.

phylogenetic level	Redundancy DNAonDOM	Redundancy RNAonDOM	Correlation & significance of 1st canonical axes DNAonDOM vs. RNAonDOM
phylum	0.30	0.28	0.89***
class	0.33	0.29	0.65*
order	0.46	0.38	0.25
family	0.44	0.38	0.34
genus	0.55	0.56	0.20
species	0.53	0.49	0.29

Significance: $p<0.05$*, $p<0.01$**, $p<0.001$***

An association between the total MC (DNA-level) and molecular DOM composition was detected from class level ($p<0.05$) and increased with lower genetic distance (Tab. 1b). The redundancy provides an estimate of the percentage of variability in the DOM matrix that can be explained by the DNA matrix, and this fraction increased with increasing resolution of the DNA matrix as well (Tab. 1c). A

significant association between the molecular DOM composition and the active community (RNA-based) was only detected at phylum level and again at genus and species resolution of the RNA-based MC. However, the first canonical axes of the CCorA of DNA/DOM and RNA/DOM are correlated at the low resolution of phylum and class (Tab. 1c), denoting that the resolution is too low to differentiate the composition of active and total community. Based on these results, we chose the "species"-level phylogenetic resolution for further in-depth analysis of possible interactions between the important contributors of the active and of the total MC and the prevailing dissolved molecules (Fig. 5).

Identifying the key players – molecular classes. In total, 6338 molecular formulae could be assigned to the samples of the reduced DOM dataset including only 12 samples. 42% of the compounds are CHO compounds, the others contain at least one heteroatom. The molecules with positive correlation to the first canonical DNA axis (n=3050, Tab. 2, Fig. 5a/b) centered in a core area above the H/C ratio of 1 which is typically observed for marine DOM and also contained some condensed aromatic compounds (Kim et al. 2003; Sleighter and Hatcher 2008). The molecules are larger (~432 Da) than the negatively correlated compounds (~351 Da). This corresponds well to the finding that terrigenous DOM is on average of smaller molecule size than the characteristic marine compounds (Koch et al. 2005). The average O/C ratio was similar for positive and negative correlations (O/C 0.46-0.48, Tab. 2). The H/C ratio of the negatively correlated molecules (n=3288, Fig. 5a) was on average 1.21, and thus marginally lower than for the positively correlated molecules, whose higher H/C ratio of ~1.27 represents the higher saturation of the molecules. The classification of the molecular formulae into molecular groups emphasizes the differences between the positively and negatively correlated molecules. The reason for this is that the H/C and O/C ratios are weighted by average peak intensity. The higher peaks are generally found in the center of the van Krevelen diagram, thus blurring the differences between positive and negative correlations. A negative correlation with the total MC was accompanied by proportions of condensed molecules (black carbon, polyphenols) that were more

than twice those of the positively correlated molecules, while the contribution of highly unsaturated compounds was lower (Tab. 2).

For RNA-DOM interrelations, the negatively correlated molecules were fewer (n=1848, Fig. 5c) and concentrated at a high H/C and O/C ratios typical for protein-like substances, but did not contain excessive amounts of nitrogen. Positively correlated molecules on the other hand were numerous (n=4490, Fig. 5c) for the active MC and widely spread over the whole van Krevelen space with a slight accumulation towards lower H/C and O/C values (Fig. 5c). The average values are again influenced by the higher intensity peaks being located towards the center of the van Krevelen space and similar to the DNA-related molecules, a size difference exists between the positively (~427 Da) and negatively (~368 Da) correlated molecules. The relative contribution of polyphenol and black carbon molecules decreased, highly unsaturated molecules were similar, and highly unsaturated aliphatics were enriched in the groups of negatively compared to positively correlating compounds (Tab. 2).

Tab. 2: Molecular characteristics and distribution of molecular classes of positively (+) and negatively (-) correlated molecules on DNA and RNA level. Values represent averages weighted by relative peak intensity, percentages are based on the total number of assigned molelcular formulae per group (DNA+, DNA-, RNA+ and RNA-).

	DNA +	DNA -	RNA +	RNA -
Number of molecules (n)	3050	3288	4490	1848
average size (Da)	432	351	427	368
O/C	0.48	0.46	0.46	0.48
H/C	1.27	1.21	1.19	1.28
% black.carbon	5.4	13.8	11.1	6.6
% polyphenols	9.9	24.8	20.5	10.6
% highly unsaturated compounds	72.9	50.5	60.3	63.7
% unsaturated aliphatics	6.2	8.9	4.4	15.3
% saturated fatty acids	0.3	0.4	0.4	0.2
% sugars	0.8	0.4	0.5	0.8
% peptides	4.5	1.2	2.8	2.9
%N containing compounds	1.8	1.1	1.7	0.9

Identifying the key players – microbial community. The 30 OTUs (DNA- and RNA-based) showing the most positive/negative correlations to the molecular DOM composition, are depicted in figure 5d. The positive correlations were $r>0.43$, the negative correlations were $r<-0.55$. Although the "species" level provided the highest correlations between DNA/RNA and DOM, a drawback of this high resolution is that it does not necessarily contain more information on the respective species due to the low number of identified, sequenced microorganisms available via the existing databases. Hence, the interpretation of the results is often compelled to remain superficial in spite of the deep sequencing effort.

The influence of salinity is weaker on the total MC than on the DOM composition of the reduced dataset: with PCoA analysis of the total MC at species level, correlation to salinity was not found before the fourth PCoA ($r=-0.74$, $p<0.01$), which explained 8% of the variability in the dataset. When examining the bacterial species on DNA-level with correlations to a subgroup of DOM molecules, they are dominated by representatives of the *Proteobacteria* and *Bacteroidetes* phyla, especially the most abundant classes of *Alphaproteobacteria* and *Flavobacteria* (Fig. 5d). Positive correlations were most abundant with *Flavobacteria* (7 orders), *Gammaproteobacteria* (3 orders) and *Alphaproteobacteria* (3 orders). Negative correlations were dominated by *Alpha*- and *Betaproteobacteria*.

PCoA analysis of the active MC at species level revealed, almost identical to the total MC, a correlation to salinity on the fourth PCoA ($r=-0.74$, $p<0.01$) which explained 11.5% of the variability in the dataset. Dominant representatives of the active MC were negatively correlated *Alphaproteobacteria* (6 orders), C*yanobacteria* (Fig. 5e). *Flavobacteria* (5 species), *Betaproteobacteria* (4 orders) and *Alphaproteobacteria* (3 orders) were found to correlate positively to the DOM molecular composition. Of the most influential representatives of the active and total MC, five species appeared in both trees (Fig. 5d/e, 2 *Flavobacteria*, 1 *Alphaproteobacterium*, 2 *Gammaproteobacteria*).

Molecule classes and microbial species – how do they relate? Opposing views exist concerning the degradation/uptake of the highly diverse organic matter by

microorganisms in the ocean: On the one hand, a taxa-specific preference could be shown by Cottrell and Kirchman (2000) who found that representatives of the *Cytophaga-Flavobacteria* cluster dominated high molecular weight organic matter degradation processes. On the other hand, many bacterioplankton species are generalists and functional redundancy of metabolic capabilities leads to a minor influence of DOM quality on MC structure (Mou et al. 2007). With our study, we could not resolve any causal effects shaping MC composition and DOM pool but provide first insights into co-occurrences (positive correlations) and absences (negative correlations) between certain molecular groups and microbes.

The canonical variates for the RNA/DOM and DNA/DOM analysis are not related, indicating that the association of the active community with DOM is not simply a subgroup but completely different from the association of the total community with DOM. However, the first canonical variate relating DOM and total MC is strongly associated with salinity (r=-0.78). So even if DOM composition and DNA-based MC are individually only weakly related to salinity, the canonical correlation between these two brings out the connection to the important driver of the system North Sea. The representation of the molecular formulae in van Krevelen space clearly shows this separation into “marine” compounds at higher H/C ratios and “terrigenous” compounds below an H/C ratio of 1 (Fig. 5a/b, (Kim et al. 2003; Koch et al. 2005; Sleighter and Hatcher 2008)). The marine DOM is positively associated with *Flavobacteria*, mainly of the *Flavobacteriales* order, and *Gammaproteobacteria*, while the terrigenous DOM is associated mainly with *Alpha-* and *Betaproteobacteria*. The presence of a general salinity gradient is in accordance with Fortunato et al. (2012), who established salinity as the main driver of MC composition along a pronounced salinity gradient from river to ocean. Also, the dominance of *Flavo-*, *Alpha-* and *Gammaproteobacteria* is in line with previous studies from the southern North Sea (Alderkamp et al. 2006; Rink et al. 2011). Bacteria of the *Bacteroidetes* phylum are known to be efficient degraders of high molecular weight organic matter (Cottrell and Kirchman 2000), the capacity to metabolize polymers and a particle-attached lifestyle are coded in the genome (Fernandez-Gomez et al. 2013). *Bacteroidetes* are abundant in most marine

environments (review by Kirchman 2002, Alonso et al. 2007) from coastal waters to the oligotrophic ocean; a preference matching the terrigenous/marine DOM gradient found in this study cannot be derived for this group. *Alphaprotebacteria*, much like *Bacteroidetes*, are abundant in many marine but also freshwater habitats, as shown for the *Rosebacter* clade in detail by Buchan et al. (2005). To some extent contrary to our findings, Rink et al. (2011) found a positive correlation for free-living *Alphaproteobacteria* with salinity in the North Sea South of 55°N in the year 2002, but could not confirm this trend one year later. The cosmopolitan *NOR5/OM60* clade of the *Gammaproteobacteria* class which account for up to 11% of a coastal bacterioplankton community during summer (Fuchs et al. 2007) has preferably been detected in nearshore surface samples (Yan et al. 2009) – contrary to our positive association of most *Gammaproteobacteria* with marine-type DOM. Testing all *Gammaproteobacteria* species in the DNA dataset to make sure this finding is not an artifact of the restriction to the 30 species with highest/lowest correlations, this trend was confirmed: 70 of the 106 species show a positive correlation with marine DOM. The association of terrigenous DOM with *Betaproteobacteria* on the other hand may reflect the riverine input, since representatives of the *Betaproteobacteria* class are more abundant in freshwater (Kirchman 2002). In the future, in order to resolve the interrelation of specific molecules and distinct metabolic pathways of certain microbes, studies in single species cultures and habitats with less pronounced environmental and hydrographic gradients need to be conducted.

Overall, the correlation between DNA and DOM is quite strong (Tab. 1b) and mainly driven by salinity; thus we hypothesize that the DOM and the total MC undergo mixing and transformation processes from the coast towards the Atlantic on similar timescales and thus retain the history of the water masses, although the involved processes may be different: e.g. the MC may respond to changes in temperature or nutrient supply (Fuhrman et al. 2008), while the DOM may be more affected by abiotic adsorption/desorption to surfaces (Gogou and Repeta 2010), flocculation (Mikes et al. 2004) or photochemical processes (Stubbins et al. 2010).

The molecules that negatively correlate to the RNA-based active MC are mainly found at high H/C and O/C ratios, and likely consist of more saturated compounds that can easily be taken up by bacteria. The microbes thriving on these compounds are *Flavo-* and *Alphaproteobacteria*, known to be metabolically diverse groups. However, the relationship could also be interpreted the other way around: a negative correlation may imply rapid uptake of the respective molecules by the respective microorganisms. Alphaproteobacteria would then be the most abundant class driving this molecule group, followed by minor contributions from *Cyano-* and *Flavobacteria* (Fig. 5e). The positively correlating molecules are too widespread over a whole range of carbon to oxygen or hydrogen ratios to draw conclusions concerning their origin or preferential use; on the MC side the occurrence of three *Cyanobacteria* within the top 15 group is noticeable, since they did not appear in the group of the active bacteria with a negative correlation to the DOM. *Prochlorococcus* for example, the most abundant genus of the *Cyanobacteria* class, has been shown to outcompete other bacterioplankton for uptake of amino acids (Zubkov et al. 2003). The general correlation between active MC and DOM was weaker than the DNA/DOM correlation (Tab. 1b). A reason for this may be the method applied for the characterization of the DOM pool: solid phase extraction of the samples with an efficiency of ~50% may not be suitable to representatively catch the most labile component of the DOM such as amino acids or sugars that is rapidly turned over by the diverse, active MC. Together with the fact that this relationship is not influenced by salinity and not correlated to the RNA/DNA canonical variate, we propose that the correlations between RNA-based MC and DOM are driven by "patchy", short term events that are not covered by our survey effort. These restricted events may include small phytoplankton blooms, or confined particle-rich water masses caused by turbulent, wind-driven mixing. Due to the multitude of possible local events in the highly dynamic North Sea on both spatial and temporal scales, a much higher sampling resolution is required to cover the natural variability of both DOM and MC composition to an extent sufficient for revealing general trends of DOM-microbe interactions.

Conclusion. The molecular composition of the DOM pool and the composition of the total as well as of the active community of the North Sea were studied in unprecedented detail in this study. Based on molecular level analysis of DOM via FT ICR MS, we found that terrigenous DOM identified as dissolved black carbon and polyphenolic compounds makes up a considerable amount of the DOM pool that can be traced from the coast, where it enters the North Sea as river/subterranean discharge, northwards to the Atlantic Ocean. The total coastal MC is transported together with the DOM as the water moves northwards and undergoes changes on a similar timescale of several months to a year. Thus, DOM and total MC can act as tracers of the intermediate water mass history. The association of the active MC to the molecular DOM composition is weak, likely because the applied methods of DOM characterization are not suitable to catch changes on the short timescales and the activity of the bacterioplankton may be triggered by local and sporadic events not exhaustively covered in this study. However, the ultrahigh resolution of the methods that were applied in this study, the elucidation of the microbial community structure as well as the molecular DOM composition together with the statistical tools applied, allow a first glimpse at the identification of the key players on both sides.

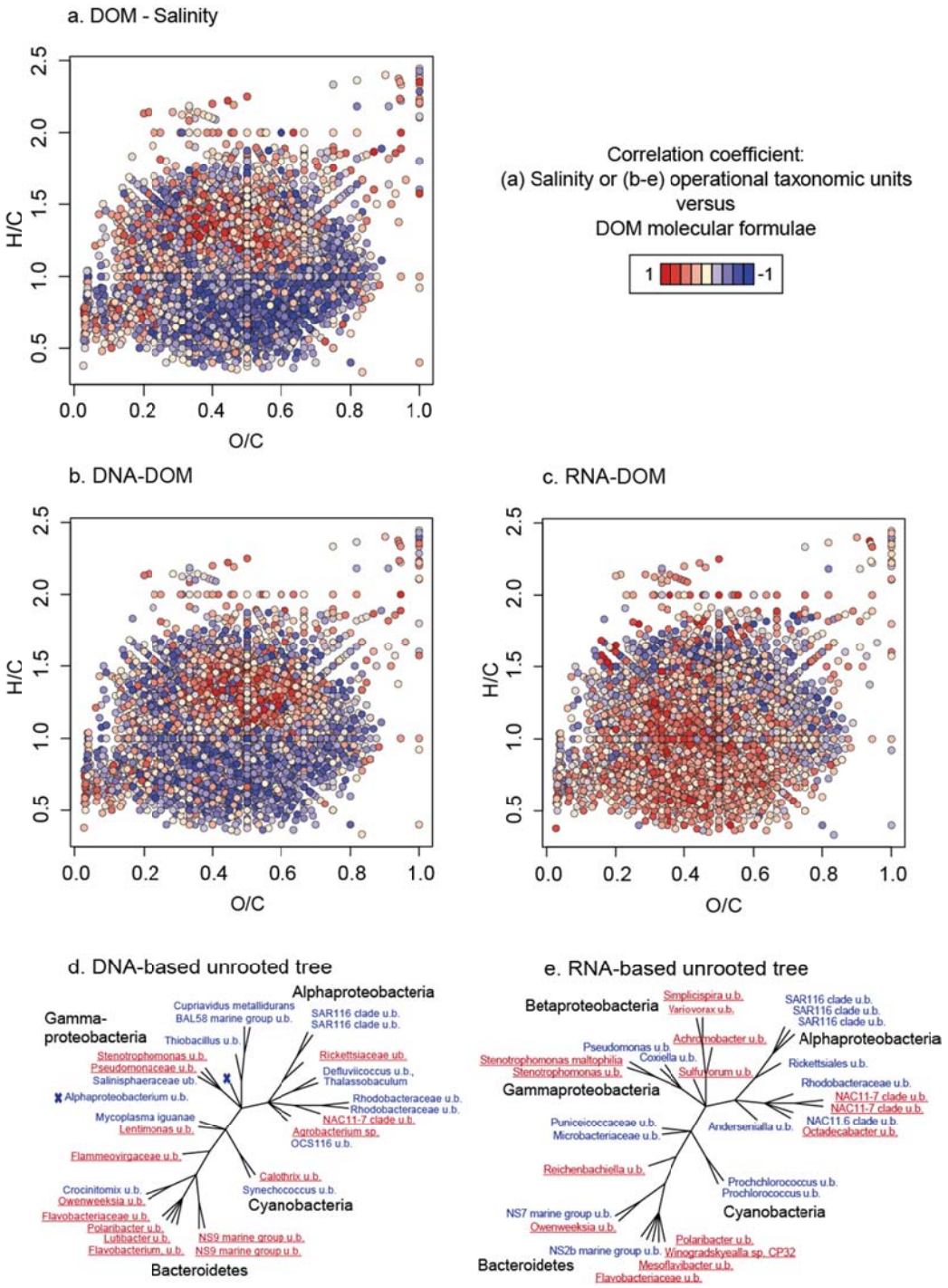

Figure 5: Correlation of molecular formulae with salinity (a); molecular formulae correlated to the canonical axis associating DNA and DOM (b) and RNA and DOM (c); Unrooted phylogenetic trees of OTUs on species level correlated to the canonical axis associating DOM and DNA (d) and DOM and RNA (e). The species with the 15 most positive (red, underlined) and the 15 most negative (blue) correlations are shown, respectively.

4.5 Acknowledgements

We thank the Captain, crew and scientific party of cruise HE361 for a successful cruise. We are grateful to Katrin Klaproth for assistance with processing FT ICR MS data and to Matthias Friebe for DOC and TDN analysis. Bernd Wemheuer and Prof. Dr. Rolf Daniel are thanked for providing the pyrosequencing data and answering numerous questions concerning data handling.

4.6 References

Abril, G., M. Nogueira, H. Etcheber, G. Cabeçadas, E. Lemaire, and M. J. Brogueira. 2002. Behaviour of Organic Carbon in Nine Contrasting European Estuaries. Estuarine, Coastal and Shelf Science **54:** 241-262.

Alderkamp, A.-C., E. Sintes, and G. J. Herndl. 2006. Abundance and activity of major groups of prokaryotic plankton in the coastal North Sea during spring and summer. Aquatic Microbial Ecology **45:** 237-246.

Alonso, C., F. Warnecke, R. Amann, and J. Pernthaler. 2007. High local and global diversity of Flavobacteria in marine plankton. Environ Microbiol **9:** 1253-1266.

Anderson, M. J., and T. J. Willis. 2003. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. Ecology **84:** 511-525.

Azam, F. 1998. Microbial Control of Oceanic Carbon Flux: The Plot Thickens. Science **280:** 694-696.

Azam, F., T. Fenchel, J. G. Field, J. S. Gray, L. A. Meyerreil, and F. Thingstad. 1983. The Ecological Role of Water-Column Microbes in the Sea. Marine Ecology-Progress Series **10:** 257-263.

Benner, R., and S. Ziegler. 2000. Do photochemical transformations of dissolved organic matter produce biorefractory as well as bioreactive substrates?, p. 181-192. Microbial biosystems: new frontiers. Proceedings of the 8th International Symposium on Microbial Ecology.

Bray, J. R., and J. T. Curtis. 1957. An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecological Monographs **27:** 326-349.

Brion, N., W. Baeyens, S. De Galan, M. Elskens, and R. P. M. Laane. 2004. The North Sea: source or sink for nitrogen and phosphorus to the Atlantic Ocean? Biogeochemistry **68:** 277-296.

Buchan, A., J. M. Gonzalez, and M. A. Moran. 2005. Overview of the marine roseobacter lineage. Appl Environ Microbiol **71:** 5665-5677.

Cadée, G. C. 1982. Tidal and seasonal variation in particulate and dissolved organic carbon in the western dutch Wadden Sea and Marsdiep tidal inlet. Netherlands Journal of Sea Research **15:** 228-249.

Campbell, B. J., L. Y. Yu, J. F. Heidelberg, and D. L. Kirchman. 2011. Activity of abundant and rare bacteria in a coastal ocean. Proc. Natl. Acad. Sci. U. S. A. **108:** 12776-12781.

Carlson, C. A., P. A. Del Giorgio, and G. J. Herndl. 2007. Microbes and the Dissipation of Energy and Respiration: From Cells to Ecosystems. Oceanography **20:** 89-100.

Colijn, F., and G. C. Cadée. 2003. Is phytoplankton growth in the Wadden Sea light or nitrogen limited? Journal of Sea Research **49:** 83-93.

Cottrell, M. T., and D. L. Kirchman. 2000. Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low- and high-molecular-weight dissolved organic matter. Appl. Environ. Microbiol. **66:** 1692-1697.

Danielssen, D. S., E. Svendsen, and M. Ostrowski. 1996. Long-term hydrographic variation in the Skagerrak based on the section Torungen–Hirtshals. ICES Journal of Marine Science: Journal du Conseil **53:** 917-925.

Dittmar, T., B. Koch, N. Hertkorn, and G. Kattner. 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. Limnology and Oceanography-Methods **6:** 230-235.

Dittmar, T., and J. Paeng. 2009. A heat-induced molecular signature in marine dissolved organic matter. Nature Geoscience **2:** 175-179.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics **26:** 2460-2461.

Edwards, R. A., and E. A. Dinsdale. 2007. Marine Environmental Genomics: Unlocking the Ocean's Secrets. Oceanography **20:** 56-61.

Eisma, D., and J. Kalf. 1987. Dispersal, concentration and deposition of suspended matter in the North Sea. Journal of the Geological Society **144:** 161-178.

Fernandez-Gomez, B. and others 2013. Ecology of marine Bacteroidetes: a comparative genomics approach. Isme Journal **7:** 1026-1037.

Fortunato, C. S., L. Herfort, P. Zuber, A. M. Baptista, and B. C. Crump. 2012. Spatial variability overwhelms seasonal patterns in bacterioplankton communities across a river to ocean gradient. Isme Journal **6:** 554-563.

Fuchs, B. M. and others 2007. Characterization of a marine gammaproteobacterium capable of aerobic anoxygenic photosynthesis. Proceedings of the National Academy of Sciences **104:** 2891-2896.

Fuhrman, J. A. and others 2008. A latitudinal diversity gradient in planktonic marine bacteria. Proceedings of the National Academy of Sciences **105:** 7774-7778.

Giebel, H. A. and others 2011. Distribution of Roseobacter RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. ISME J **5:** 8-19.

Gogou, A., and D. J. Repeta. 2010. Particulate-dissolved transformations as a sink for semi-labile dissolved organic matter: Chemical characterization of high molecular weight dissolved and surface-active organic matter in seawater and in diatom cultures. Mar. Chem. **121:** 215-223.

Goldberg, E. D. 1985. Black carbon in the environment. Properties and distribution. John Wiley & Sons Inc.

Guo, L. D., J. Z. Zhang, and C. Gueguen. 2004. Speciation and fluxes of nutrients (N, P, Si) from the upper Yukon River. Global Biogeochemical Cycles **18**.

Hansell, D. A., and C. A. Carlson. 1998. Deep-ocean gradients in the concentration of dissolved organic carbon. Nature **395:** 263-266.

Hedges, J. I. 1992. Global biogeochemical cycles - progress and problems. Mar. Chem. **39:** 67-93.

Herndl, G. J., and T. Reinthaler. 2013. Microbial control of the dark end of the biological pump. Nature Geosci **6:** 718-724.

Hernes, P. J., and R. Benner. 2002. Transport and diagenesis of dissolved and particulate terrigenous organic matter in the North Pacific Ocean. Deep Sea Research Part I: Oceanographic Research Papers **49:** 2119-2132.

Hernes, P.J., 2006. Terrigenous organic matter sources and reactivity in the North Atlantic Ocean and a comparison to the Arctic and Pacific oceans. Mar. Chem. **100:** 66-79.

Hutchinson, G. E. 1961. The paradox of the plankton. American Naturalist **95:** 137-145.

Jaffé, R. and others 2013. Global Charcoal Mobilization from Soils via Dissolution and Riverine Transport to the Oceans. Science **340:** 345-347.

Jiao, N. and others 2010. Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. Nat Rev Microbiol **8:** 593-599.

Jones, S. E., and J. T. Lennon. 2010. Dormancy contributes to the maintenance of microbial diversity. Proc. Natl. Acad. Sci. U. S. A. **107:** 5881-5886.

Kerner, M., and A. Spitzy. 2001. Nitrate Regeneration Coupled to Degradation to Different Size Fractions of DON by the Picoplankton in the Elbe Estuary. Microbial Ecology **41:** 69-81.

Kim, S., R. W. Kramer, and P. G. Hatcher. 2003. Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram. Analytical Chemistry **75:** 5336-5344.

Kirchman, D. L. 2002. The ecology of Cytophaga-Flavobacteria in aquatic environments. Fems Microbiology Ecology **39:** 91-100.

Koch, B. P., and T. Dittmar. 2006. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. Rapid Commun. Mass Spectrom. **20:** 926-932.

Koch, B. P., M. R. Witt, R. Engbrodt, T. Dittmar, and G. Kattner. 2005. Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Geochimica Et Cosmochimica Acta **69:** 3299-3308.

Kraft, Nathan j. B., William k. Cornwell, Campbell o. Webb, and David d. Ackerly. 2007. Trait Evolution, Community Assembly, and the Phylogenetic Structure of Ecological Communities. The American Naturalist **170:** 271-283.

Kujawinski, E. B. 2002. Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS): characterization of complex environmental mixtures. Environ. Forensics **3:** 207-216.

Legendre, P., and L. Legendre. 1998. Numerical Ecology, Second Edition ed. Elsevier Science.

Mikes, D., R. Verney, R. Lafite, and M. Belorgey. 2004. Controlling factors in estuarine flocculation processes: experimental results with material from the Seine Estuary, Northwestern France. Journal of Coastal Research**:** 82-89.

Miller, W. L., and M. A. Moran. 1997. Interaction of photochemical and microbial processes in the degradation of refractory dissolved organic matter from a coastal marine environment. Limnology and Oceanography **42:** 1317-1324.

Moran, M. A., and R. E. Hodson. 1994. Support of Bacterioplankton Production by Dissolved Humic Substances from 3 Marine Environments. Marine Ecology Progress Series **110:** 241-247.

Mou, X. Z., R. E. Hodson, and M. A. Moran. 2007. Bacterioplankton assemblages transforming dissolved organic compounds in coastal seawater. Environmental Microbiology **9:** 2025-2037.

Nixon, S. W. and others 1996. The Fate of Nitrogen and Phosphorus at the Land-Sea Margin of the North Atlantic Ocean. Biogeochemistry **35:** 141-180.

Oksanen, J. and others 2012. vegan: Community Ecology Package. R package version 2.0-4. http://CRAN.R-project.org/package=vegan.

Opsahl, S., and R. Benner. 1997. Distribution and cycling of terrigenous dissolved organic matter in the ocean. Nature **386:** 480-482.

Opsahl, S., R. Benner, and R. M. W. Amon. 1999. Major flux of terrigenous dissolved organic matter through the Arctic Ocean. Limnology and Oceanography **44:** 2017-2023.

Radach, G., and J. Pätsch. 1997. Climatological annual cycles of nutrients and chlorophyll in the North Sea. Journal of Sea Research **38:** 231-248.

Radach, G., and J. Pätsch., 2007. Variability of continental riverine freshwater and nutrient inputs into the North Sea for the years 1977–2000 and its consequences for the assessment of eutrophication. Estuaries and Coasts **30:** 66-81.

Riedel, T., D. Zak, H. Biester, and T. Dittmar. 2013. Iron traps terrestrially derived dissolved organic matter at redox interfaces. Proceedings of the National Academy of Sciences **110:** 10101-10105.

Rink, B., N. Grüner, T. Brinkhoff, K. Ziegelmüller, and M. Simon. 2011. Regional patterns of bacterial community composition and biogeochemical properties in the southern North Sea. Aquatic Microbial Ecology **63:** 207-222.

Šantl-Temkiv, T. and others 2013. Hailstones: A Window into the Microbial and Chemical Inventory of a Storm Cloud. PLoS ONE **8:** e53550.

Shaffer, G. 1996. Biogeochemical cycling in the global ocean .2. New production, Redfield ratios, and remineralization in the organic pump. J. Geophys. Res.-Oceans **101:** 3723-3745.

Sleighter, R. L., and P. G. Hatcher. 2008. Molecular characterization of dissolved organic matter (DOM) along a river to ocean transect of the lower Chesapeake Bay by ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Mar. Chem. **110:** 140-152.

Stegen, J. C., X. Lin, A. E. Konopka, and J. K. Fredrickson. 2012. Stochastic and deterministic assembly processes in subsurface microbial communities. ISME J **6:** 1653-1664.

Stubbins, A. and others 2010. Illuminated darkness: Molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry. Limnology and Oceanography **55:** 1467-1477.

Suratman, S., K. Weston, T. Jickells, and L. Fernand. 2009. Spatial and seasonal changes of dissolved and particulate organic C in the North Sea. Hydrobiologia **628:** 13-25.

Teeling, H. and others 2012. Substrate-Controlled Succession of Marine Bacterioplankton Populations Induced by a Phytoplankton Bloom. Science **336:** 608-611.

Thomas, H. and others 2005. The carbon budget of the North Sea. Biogeosciences **2:** 87-96.

Thomas, H., J. P. Gattuso, and S. V. Smith. 2003. Coastal biogeochemistry. EGS-AGU-EUG Joint Assembly. LOICZ Newsletter.

Tranvik, L. J. 1992. Allochthonous dissolved organic matter as an energy source for pelagic bacteria and the concept of the microbial loop. Hydrobiologia **229:** 107-114.

Turrell, W. R. 1992. New hypotheses concerning the circulation of the northern North Sea and its relation to North Sea fish stock recruitment. ICES Journal of Marine Science: Journal du Conseil **49:** 107-123.

Valentin-Vargas, A., G. Toro-Labrador, and A. A. Massol-Deya. 2012. Bacterial Community Dynamics in Full-Scale Activated Sludge Bioreactors: Operational and Ecological Factors Driving Community Assembly and Performance. PLoS ONE **7**.

Wetzel, R. G. 1984. Detrital Dissolved and Particulate Organic Carbon Functions in Aquatic Ecosystems. Bulletin of Marine Science **35:** 503-509.

Yan, S. and others 2009. Biogeography and phylogeny of the NOR5/OM60 clade of Gammaproteobacteria. Syst Appl Microbiol **32:** 124-139.

Zinger, L. and others 2011. Global Patterns of Bacterial Beta-Diversity in Seafloor and Seawater Ecosystems. PLoS ONE **6**.

Zubkov, M. V., B. M. Fuchs, G. A. Tarran, P. H. Burkill, and R. Amann. 2003. High rate of uptake of organic nitrogen compounds by Prochlorococcus cyanobacteria as a key to their dominance in oligotrophic oceanic waters. Appl Environ Microbiol **69:** 1299-1304.

4.7 Supplementary Results

Table 1: Position (latitude, longitude in decimal format) of stations sampled during HE361 cruise. Sampled depth, temperature and salinity were obtained by CTD sonde, dissolved organic carbon (DOC) and total dissolved nitrogen (TDN) were measured via high temperature catalytic oxidation in the laboratory. Standard deviation (s.d.) given for triplicate analyses.

Station	Latitude (degrees N)	Longitude (degrees E)	Depth (m)	Temperature (°C)	Salinity (psu)	DOC (±s.d., µmol L^{-1})	TDN (±s.d., µmol L^{-1})
1	54.0893	7.9327	3	18.3	29.83	169±1	20±1
2	54.9667	7.6058	3	15.8	32.25	126±3	14±1
			10	15.8	32.25	122±1	14±1
3	55.4662	7.2852	3	14.8	33.42	105±0	10±1
			10	14.7	33.42	107±1	10±0
5	56.0118	7.6902	3	15.8	32.35	129±1	15±1
			10	15.7	32.64	122±1	14±0
6	56.5018	7.5015	3	15.7	33.22	110±3	11±1
			10	15.7	33.25	104±1	9±1
			23	11.8	34.24	95±6	8±0
7	56.9940	7.2925	3	15.2	33.11	98±2	7±1
			10	15.2	33.11	98±0	6±0
			27	10.1	34.68	79±1	7±1
8	57.2610	8.7320	3	14.8	33.41	101±0	7±0
			10	14.8	33.41	97±2	7±0
			18	13.8	33.60	98±0	7±0
10	57.3173	9.0008	3	15.3	33.17	111±2	4±0
			8	15.3	33.19	108±2	3±0
			17	14.4	33.39	103±0	4±0
12	57.4887	7.1000	3	15.5	31.87	128±2	5±1
			25	9.7	34.99	79±1	2±0
			150	7.2	35.24	71±1	13±1
13	57.8305	6.9185	3	16.5	30.06	149±1	6±1
			25	12.0	33.33	93±1	3±1
			350	6.0	35.16	72±3	9±0
14	58.1657	5.1693	4	14.7	32.57	109±0	5±1
			250	6.1	35.14	84±1	12±0
15	58.9985	3.9232	3	14.6	31.86	106±1	8±1
			18	8.5	34.46	86±1	9±0
			250	6.5	35.22	71±6	15±1
16	59.9848	3.0360	4	14.5	32.17	98±1	7±0

Station	Latitude (degrees N)	Longitude (degrees E)	Depth (m)	Temperature (°C)	Salinity (psu)	DOC (±s.d., µmol L^{-1})	TDN (±s.d., µmol L^{-1})
			22	12.0	35.06	75±2	6±1
			100	9.3	35.40	55±2	15±0
27	57.6128	8.5818	3	16.0	34.08	87±1	6±0
			10	16.0	34.24	85±1	6±0
			29	12.5	34.80	78±3	6±1
			100	8.2	35.02	70±1	9±0
29	57.6115	8.5935	3	16.0	34.37	84±1	5±0
			30	12.2	34.72	81±1	6±1
33	56.0110	7.6918	3	16.7	32.97	107±2	8±0
			16	15.7	33.21	100±2	9±0
34	55.4692	7.2848	3	16.1	33.04	104±4	9±0
			20	15.5	33.32	98±1	9±1
35	54.9662	7.6032	3	17.6	32.12	118±4	16±0
			21	16.0	32.38	114±1	13±0

Table 2: Environmental parameters obtained from HE361 cruise (THDAA total hydrolysable dissolved amino acids, FDNS free dissolved neutral sugars, n.a. not analyzed).

Station	Depth (m)	Relative fluorescence	Chlorophyll *a* (µg L^{-1})	Phaeopigments (µg L^{-1})	Bacterial cell counts (cells mL^{-1} $\times 10^6$)	Biomass production (nmol L^{-1} h^{-1})	THDAA [nmol L^{-1}]	FDNS [µmol L^{-1}]
1	3	0.76	2.01	0.79	1.66	712.7	966	1.9
2	3	0.62	1.30	0.49	1.60	321.6	612	0.8
	10	0.61	1.21	0.43	1.71	381.3	638	-1
3	3	0.38	0.53	0.38	0.76	155.2	588	0
	10	0.44	0.69	0.35	0.80	181.8	606	0
5	3	0.85	1.34	0.43	0.87	156.6	813	0.1
	10	1.00	1.22	0.61	1.04	296.6	n.a.	1.0
6	3	0.39	0.79	0.25	0.97	420.0	716	0.7
	10	0.54	0.91	0.27	0.93	391.6	n.a.	0.2
	23	2.14	2.50	0.26	0.73	198.8	447	0.6
7	3	0.39	0.28	0.19	0.86	149.4	723	0.4
	10	0.47	0.39	0.12	0.95	139.2	n.a.	0.0
	27	0.75	0.87	0.17	1.22	94.6	381	0.1
8	3	1.49	1.60	0.61	0.72	252.6	548	0.8
	10	1.60	1.99	0.42	0.77	202.6	542	0.4
	18	1.73	2.70	0.80	0.84	n.a.	n.a.	0.5
10	3	1.54	3.16	0.46	0.67	316.4	589	0.3
	8	2.18	2.49	0.73	0.69	302.7	492	0

Station	Depth (m)	Relative fluorescence	Chlorophyll *a* (µg L^{-1})	Phaeopigments (µg L^{-1})	Bacterial cell counts (cells mL^{-1} $\times 10^6$)	Biomass production (nmol L^{-1} h^{-1})	THDAA [nmol L^{-1}]	FDNS [µmol L^{-1}]
	17	1.79	2.80	0.69	0.69	257.7	591	0
12	3	0.36	0.22	0.13	0.58	161.8	604	0.7
	25	1.04	1.28	0.23	2.56	82.4	367	0
	150	0.18	n.a.	n.a.	n.a.	n.a.	623	0.3
13	3	0.50	0.47	0.03	0.82	147.0	636	1.9
	25	1.59	-0.47	1.34	1.79	106.7	n.a.	0.1
	350	0.20	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
14	4	0.57	0.61	0.07	1.10	108.3	799	0.2
	250	0.17	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
15	3	0.40	0.30	0.10	0.91	183.6	529	1.0
	18	0.94	0.93	-0.04	1.37	89.1	352	0.1
	250	0.18	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
16	4	0.47	0.46	0.04	1.14	196.8	414	0.4
	22	2.25	1.61	0.47	2.61	115.5	n.a.	0.1
	100	0.16	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
27	3	0.50	0.39	0.01	0.79	179.2	456	0.3
	10	0.58	0.59	-0.15	0.65	143.6	n.a.	0.0
	29	0.98	1.22	0.27	n.a.	n.a.	405	0
	100	0.54	n.a.	n.a.	n.a.	n.a.	n.a.	1.4
29	3	0.25	0.28	0.11	0.68	155.0	418	0
	30	1.09	1.47	0.18	n.a.	n.a.	422	1.7
33	3	0.56	1.03	0.32	0.96	500.9	620	-1
	16	1.26	3.26	0.89	0.89	391.3	569	0.6
34	3	0.28	0.41	0.20	0.61	341.3	542	0.2
	20	1.21	2.78	1.02	0.83	n.a.	n.a.	0.6
35	3	0.51	0.87	-0.06	1.11	397.9	617	0.8
	21	1.38	2.37	0.84	1.24	455.0	n.a.	n.a.

5 Concluding remarks and perspectives

This thesis investigated microbial interactions with DOM in a variety of settings. The production and cycling of DOM on molecular level was in the focus of comprehensive studies including environmental gradients and microbial communities. Additionally, the links between molecular DOM composition and associated reactivities of the compounds were emphasized. Below, the thesis is summarized into three broad conclusions, along with a short discussion about their implications and some potential future research.

Ultrahigh resolution mass spectrometry provides novel insights into marine DOM composition. The application of ESI FT ICR MS in natural organic matter research has picked up since the first publications by Stenson et al. in 2002 and Kujawinski in 2002. Koch et al. published the first molecular-level characterization of marine DOM in 2005. In this thesis, ultrahigh resolution mass spectrometry was successfully applied to several purposes: The molecular DOM composition in the high-latitude fjords of the Svalbard archipelago was explored for the first time. In the North Sea, terrigenous compounds could be traced due to their distinct H/C and O/C ratios. Furthermore, the technique proved to be irreplaceable in the discrimination of truly refractory from semi-refractory or semi-labile DOM in laboratory batch incubations. It has, however, been shown that complementary analyses of e.g. amino acids or carbohydrates are very valuable for the understanding of the total carbon pool dynamics, since these compounds are not comprehensively covered by FT ICR MS analysis due to the constraints of the applied method including sample processing (Herlemann et al. 2014). The analytical window is determined first of all by the extraction method, which is necessary to concentrate and desalt the DOM before the sensitive mass spectrometric analysis. The frequently applied PPL extraction (Dittmar et al. 2008) non-selectively adsorbs a wide range of polar to apolar compounds, but some C enrichment/N depletion occurs (Green et al. 2014) with extraction efficiencies of ~60% for oceanic DOM. Overall, 20 to 80% of the original DOM can be extracted by current standard

methods solid phase extraction, reverse osmosis/electrodialysis or ultrafiltration (Benner et al. 1997; Green et al. 2014). Thus, the most suitable method needs to be chosen carefully for each study to achieve the most comprehensive sample for the characterization of the oceanic, but also the limnic or terrigenous DOM pools. Another constraint of the ultrahigh resolution mass spectrometric analysis is the electrospray ionization, which is capable of ionizing intact molecules, but the ionization mechanism of mixtures is not very well understood and the ionization of complex mixtures can hardly be performed quantitatively (Koivusalo et al. 2001). Polyethylene glycol was used as an internal standard for a better comparison between FT ICR spectra in this thesis (Chapter 3). These first results were promising and further investigation of matrix effects, the impact on the ionization, but also the testing of other potential standard compounds could significantly improve the comparison of spectra between sample sets towards "quantitative" FT ICR MS analyses. An essential drawback of DOM analysis via FT ICR MS is the lack of structural information. The fragmentation of all compounds resolved on one nominal mass in the FT ICR MS revealed that the same molecular formulae detected in DOM from constrained laboratory incubations and deep ocean DOM represent molecules of different structural composition (Chapter 3). The structure of a compound can have significant influence on its reactivity (Drastik et al. 2013; Shimizu et al. 2012) and change its susceptibility to biotic or abiotic degradation. The extension of studies elucidating the variation of structural composition and diversity of DOM, e.g. via FT ICR MS fragmentation experiments or NMR, between different environments could lead to a better understanding of the hidden DOM reactivities.

Biological processes rapidly transform labile substrates and produce highly diverse DOM. The production of highly diverse DOM by a microbial community was observed on molecular level for the first time and it could be shown that deep ocean refractory DOM can be produced by biological processes alone (Chapter 3). This refractory DOM with the exact same FT ICR MS intensity distribution pattern of molecules has been detected in Svalbard fjords (Chapter 2) as well as in other distant oceanic realms including the Antarctic (Lechtenfeld et al. 2014) or the Pacific Ocean (Hertkorn et al. 2006). In accordance with the reactivity continuum

described by Hansell (2013), the huge refractory DOM pool is similar in molecular composition and ubiquitously present in the world's oceans whereas the small labile, semi-labile and semi-refractory DOM pools carry the imprint of recent processes such as phytoplankton production, microbial reworking or abiotic reactions. This small fraction sufficed for the establishment of correlations between the total as well as the active microbial community to the molecular DOM signature (Chapter 4).

Independent of the respective microbial community composition, DOM of a highly consistent molecular DOM signature was produced in experimental incubations (Chapter 2). This finding is most likely a product of the diverse, yet habitat and resource adapted, metabolic capabilities of the microbial community (Delong et al. 2006) that occupies all existing ecological niches (Jaspers et al. 2001), supported by the global ocean circulation. The average age of refractory DOM is estimated at 16,000 years (Hansell 2013) and exceeds the ocean's turnover time of less than 1200 years (Matsumoto 2007) along the conveyor belt several times. However, the question remains why bacteria produce such theoretically energy- and nutrient-rich byproducts that can apparently not be utilized further (Danchin and Sekowska 2014). In this context, another implication demands further investigation: There is strong evidence that the amount of refractory DOM that is produced by the microbial community is a function of the net community production (NCP, Chapter 2). Exploring this hypothesis was outside the scope of this thesis, but the mechanisms behind it need to be untangled to infer consequences for the ocean in the face of climate change. The impact of climate change on marine DOM dynamics are yet unknown: how will global warming, ocean acidification, rising temperatures and deoxygenation (Gruber 2011) affect these cycles? The sea surface temperatures have increased by 0.7°C over the last century (Levitus et al. 2000) leading to a more pronounced stratification of the upper ocean, which hinders exchange processes for e.g. oxygen, nutrients, but also organic matter, with deeper water layers. Ocean acidification, the change in the ocean carbonate system through the uptake of anthropogenic CO_2, will have an impact on all levels of the microbial food web, and some oceanic regions such as the cold waters of the high Arctic will be impacted more heavily than others (Riebesell et al. 2013). Engel et al. (2004) were able to

show an enhanced production of transparent exopolymeric particles (TEP) by the alga *Emiliania huxley* under increased CO_2 conditions, while no increase in DOC concentration could be proven. This may be caused by an increase in bacterial degradation activity at high CO_2 level shown by Piontek et al. (2010). The faster microbial polysaccharide turnover might, in the long term, reduce the carbon export to the deep ocean and affect the CO_2 equilibrium between atmosphere and ocean. An increased CO_2 level in the surface may increase the net community production since more CO_2 is available as substrate for the primary producers, and a subsequent increase of the production of refractory DOM can in turn influence the marine carbon cycle. The relationship of the NCP and the production of refractory DOM should be explored in constrained laboratory incubation experiments with varying nutrient and CO_2 supply to different species or a mixed community of primary producers, possibly providing insights into the underlying mechanisms.

The molecular DOM signature can be linked to the microbial community composition. A relationship between phytoplankton species and molecular DOM composition is often assumed and has been demonstrated in single species cultures (Becker et al. 2014). In the partly stratified fjords of the Svalbard archipelago, however, the microbial community structure derived from DGGE fingerprinting was distinctly different between surface and bottom water, while the molecular DOM composition did not show any correlation to sampling sites and depths (Chapter 2). For this study, it could not be precluded that unknown labile, low-molecular weight or colloidal compounds that escape the analytical window (Herlemann et al. 2014) link bacterial and DOM, but a major constraint was apparently the low resolution of the DGGE fingerprinting technique compared to ultrahigh resolution mass spectrometry for molecular DOM characterization. A linkage of molecular DOM composition to the microbial community composition of the North Sea (Chapter 4) that was analyzed via the next-generation 454 deep sequencing approach yielded much better results and certain coherences could be established. A novel application of statistical tools to deal with the combination of two large datasets was described and constitutes a valuable starting point for the analysis of datasets including a higher number of samples to validate the proposed method. A factor that influences

the microbial community as well as the DOM composition is often not addressed in the examination of biogeochemical cycles: A smaller, but likely still significant, influence on the cycling of carbon compared to microbes is exerted by viruses through the "viral shunt" (Wilhelm and Suttle 1999). Rough estimates indicate that 10 to 20% of the prokaryotes are lysed daily by viruses (Suttle 1994), and their cell wall components and dissolved contents are released. Few studies have explored the bulk DOC (Bratbak et al. 1992), amino acid (Middelboe and Jørgensen 2006) and carbohydrate contribution (Weinbauer and Peduzzi 1995) or the optical signature (Lonborg et al. 2013) of DOM associated with viral lysis. So far molecular analyses and a quantitative integration in the carbon cycle are lacking and may be considered in future work.

In conclusion, the results of the studies included in this thesis illustrate that DOM composition and dynamics are irrevocably connected with the environment - abiotic, but especially the biotic processes shape the oceanic DOM pool. Interdisciplinary collaborations among scientists are crucial to advance our understanding of biogeochemical cycles on global scales. As it is time- and cost-intensive to obtain the data, a public data repository including DOM molecular composition, compared to PANGAEA (Grobe et al. 2006), would be highly useful. At the same time, the development of tools for the handling and analysis of complex datasets needs to be further emphasized.

6 References

Alldredge, A. L., U. Passow, and B. E. Logan. 1993. The abundance and significance of a class of large, transparent organic particles in the ocean. Deep Sea Research Part I: Oceanographic Research Papers **40:** 1131-1140.

Amon, R. M. W., and R. Benner. 1996. Bacterial utilization of different size classes of dissolved organic matter. Limnology and Oceanography **41:** 41-51.

Amon, R. M. W., and R. Benner. 2003. Combined neutral sugars as indicators of the diagenetic state of dissolved organic matter in the Arctic Ocean. Deep Sea Research Part I: Oceanographic Research Papers **50:** 151-169.

Arnosti, C., A. D. Steen, K. Ziervogel, S. Ghobrial, and W. H. Jeffrey. 2011. Latitudinal Gradients in Degradation of Marine Dissolved Organic Carbon. PLoS ONE **6:** e28900.

Azam, F., T. Fenchel, J. G. Field, J. S. Gray, L. A. Meyerreil, and F. Thingstad. 1983. The Ecological Role of Water-Column Microbes in the Sea. Marine Ecology-Progress Series **10:** 257-263.

Baines, S. B., and M. L. Pace. 1991. The production of dissolved organic matter by phytoplankton and its importance to bacteria: patterns across marine and freshwater systems. Limnol. Oceanogr. **36**.

Bakker, M. G., Z. J. Tu, J. M. Bradeen, and L. L. Kinkel. 2012. Implications of Pyrosequencing Error Correction for Biological Data Interpretation. PLoS ONE **7:** e44357.

Bauer, J. E., P. M. Williams, and E. R. M. Druffel. 1992. 14C activity of dissolved organic carbon fractions in the north-central Pacific and Sargasso Sea. Nature **357:** 667-670.

Becker, J. W. and others 2014. Closely related phytoplankton species produce similar suites of dissolved organic matter. Frontiers in Microbiology **5**.

Benner, R., B. Biddanda, B. Black, and M. Mccarthy. 1997. Abundance, size distribution, and stable carbon and nitrogen isotopic compositions of marine

organic matter isolated by tangential-flow ultrafiltration. Mar. Chem. **57:** 243-263.

Benner, R., P. Louchouarn, and R. M. W. Amon. 2005. Terrigenous dissolved organic matter in the Arctic Ocean and its transport to surface and deep waters of the North Atlantic. Global Biogeochemical Cycles **19:** -.

Benner, R., J. D. Pakulski, M. Mccarthy, J. I. Hedges, and P. G. Hatcher. 1992. Bulk Chemical Characteristics of Dissolved Organic Matter in the Ocean. Science **255:** 1561-1564.

Bratbak, G., M. Heldal, T. F. Thingstad, B. Riemann, and O. H. Haslund. 1992. Incorporation of viruses into the budget of microbial C-transer – a 1st approach. Marine Ecology Progress Series **83:** 273-280.

Carlson, C. A., H. W. Ducklow, D. A. Hansell, and W. O. Smith. 1998. Organic carbon partitioning during spring phytoplankton blooms in the Ross Sea polynya and the Sargasso Sea. Limnology and Oceanography **43:** 375-386.

Carr, M.-E. and others 2006. A comparison of global estimates of marine primary production from ocean color. Deep-Sea Research Part Ii-Topical Studies in Oceanography **53:** 741-770.

Cauwet, G. 2002. DOM in the Coastal Zone, p. 579-609. *In* D. A. Hansell and C. A. Carlson [eds.], Biogeochemistry of Marine Dissolved Organic Matter. Academic Press.

Cottrell, M. T., D. N. Wood, L. Y. Yu, and D. L. Kirchman. 2000. Selected chitinase genes in cultured and uncultured marine bacteria in the alpha- and gamma-subclasses of the proteobacteria. Appl. Environ. Microbiol. **66:** 1195-1201.

Crump, B. C., C. S. Hopkinson, M. L. Sogin, and J. E. Hobbie. 2004. Microbial biogeography along an estuarine salinity gradient: combined influences of bacterial growth and residence time. Appl Environ Microbiol **70:** 1494-1505.

Danchin, A., and A. Sekowska. 2014. The logic of metabolism and its fuzzy consequences. Environmental Microbiology **16:** 19-28.

Dauwe, B., and J. J. Middelburg. 1998. Amino acids and hexosamines as indicators of organic matter degradation state in North Sea sediments. Limnology and Oceanography **43:** 782-798.

Delong, E. F. and others 2006. Community genomics among stratified microbial assemblages in the ocean's interior. Science **311:** 496-503.

Dittmar, T. 2014. Reasons behind long-term stability of dissolved organic matter. *In* D. A. Hansell and C. A. Carlson [eds.], The Biogeochemistry of Marine Dissolved Organic Matter.

Dittmar, T. and others 2012. Continuous flux of dissolved black carbon from a vanished tropical forest biome. Nature Geosci **5:** 618-622.

Dittmar, T., B. Koch, N. Hertkorn, and G. Kattner. 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. Limnology and Oceanography-Methods **6:** 230-235.

Dittmar, T., and J. Paeng. 2009. A heat-induced molecular signature in marine dissolved organic matter. Nature Geoscience **2:** 175-179.

Dittmar, T., and A. Stubbins. 2014. 12.6 - Dissolved Organic Matter in Aquatic Systems, p. 125-156. *In* H. D. Holland and K. K. Turekian [eds.], Treatise on Geochemistry (Second Edition). Elsevier.

Dittmar, T., K. Whitehead, E. C. Minor, and B. P. Koch. 2007. Tracing terrigenous dissolved organic matter and its photochemical decay in the ocean by using liquid chromatography/mass spectrometry. Mar. Chem. **107:** 378-387.

Drastik, M., F. Novak, and J. Kucerik. 2013. Origin of heat-induced structural changes in dissolved organic matter. Chemosphere **90:** 789-795.

Ducklow, H. W., and C. A. Carlson. 1992. Oceanic Bacterial Production, p. 113-181. *In* K. C. Marshall [ed.], Advances in Microbial Ecology. Advances in Microbial Ecology. Springer US.

Elahi, E., and M. Ronaghi. 2004. Pyrosequencing, p. 211-219. *In* S. Zhao and M. Stodolsky [eds.], Bacterial Artificial Chromosomes. Methods in Molecular Biology™. Humana Press.

Engel, A. and others 2004. Transparent exopolymer particles and dissolved organic carbon production by Emiliania huxleyi exposed to different CO2 concentrations: a mesocosm experiment. Aquatic Microbial Ecology **34:** 93-104.

Flerus, R. and others 2012. A molecular perspective on the ageing of marine dissolved organic matter. Biogeosciences **9:** 1935-1955.

Fogg, G. E. 1983. The Ecological Significance of Extracellular Products of Phytoplankton Photosynthesis, p. 3. Botanica Marina.

Fortunato, C. S., L. Herfort, P. Zuber, A. M. Baptista, and B. C. Crump. 2012. Spatial variability overwhelms seasonal patterns in bacterioplankton communities across a river to ocean gradient. Isme Journal **6:** 554-563.

Fuhrman, J. A. and others 2008. A latitudinal diversity gradient in planktonic marine bacteria. Proceedings of the National Academy of Sciences **105:** 7774-7778.

Galand, P. E., E. O. Casamayor, D. L. Kirchman, and C. Lovejoy. 2009. Ecology of the rare microbial biosphere of the Arctic Ocean. Proceedings of the National Academy of Sciences **106:** 22427-22432.

Geller, A. 1985. Degradation and formation of refractory DOM by bacteria during simultaneous growth on labile substrates and persistent lake water constituents. Schweiz. Z. Hydrol **47:** 27-44.

Green, N. W. and others 2014. An intercomparison of three methods for the large-scale isolation of oceanic dissolved organic matter. Mar. Chem.

Grobe, H., M. Diepenbroek, N. Dittert, M. Reinke, and R. Sieger. 2006. Archiving and distributing earth-science data with the PANGAEA information system. *In* H. M. Dieter Fütterer; Detlef Damaske; Georg Kleinschmidt, Franz Tessensohn [ed.], Proceedings of the IX International Symposium of Antarctic Earth Sciences 2003. Springer.

Grossart, H. P. 2011. Ecological consequences of bacterioplankton lifestyles: changes in concepts are needed. Environ. Microbiol. Rep. **2:** 706-714.

Gruber, N. 2011. Warming up, turning sour, losing breath: ocean biogeochemistry under global change. Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci. **369:** 1980-1996.

Han, J. and others 2008. Towards high-throughput metabolomics using ultrahigh-field Fourier transform ion cyclotron resonance mass spectrometry. Metabolomics **4:** 128-140.

Hansell, D. A. 2013. Recalcitrant Dissolved Organic Carbon Fractions. Annual Review of Marine Science **5:** 421-445.

Hansell, D. A., C. A. Carlson, D. J. Repeta, and R. Schlitzer. 2009. Dissolved organic matter in the ocean - a controversy stimulates new insights. Oceanography **22:** 202-211.

Hedges, J. I. 1992. Global biogeochemical cycles - progress and problems. Mar. Chem. **39:** 67-93.

Hedges, J. I. and others 1994. Origins and Processing of Organic Matter in the Amazon River as Indicated by Carbohydrates and Amino Acids. Limnology and Oceanography **39:** 743-761.

Hedges, J. I., R. G. Keil, and R. Benner. 1997. What happens to terrestrial organic matter in the ocean? Organic Geochemistry **27:** 195-212.

Herlemann, D. P. R. and others 2014. Uncoupling of bacterial and terrigenous dissolved organic matter dynamics in decomposition experiments. PLoS ONE **9:** e93945-e93945.

Herndl, G. J., and T. Reinthaler. 2013. Microbial control of the dark end of the biological pump. Nature Geosci **6:** 718-724.

Hernes, P. J., and R. Benner. 2006. Terrigenous organic matter sources and reactivity in the North Atlantic Ocean and a comparison to the Arctic and Pacific oceans. Mar. Chem. **100:** 66-79.

Hertkorn, N. and others 2006. Characterization of a major refractory component of marine dissolved organic matter. Geochimica Et Cosmochimica Acta **70:** 2990-3010.

Hertkorn, N., M. Harir, B. P. Koch, B. Michalke, and P. Schmitt-Kopplin. 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.

Hertkorn, N. and others 2007. High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems. Anal Bioanal Chem **389:** 1311-1327.

Hopkinson, C. S., B. Fry, and A. L. Nolin. 1997. Stoichiometry of dissolved organic matter dynamics on the continental shelf of the northeastern U.S.A. Continental Shelf Research **17:** 473-489.

Hugenholtz, P. 2002. Exploring prokaryotic diversity in the genomic era. Genome Biol **3:** 29.

Hunt, D. E., E. Ortega-Retuerta, and C. E. Nelson. 2010. Connections between bacteria and organic matter in aquatic ecosystems: Linking microscale ecology to global carbon cycling. Eco-DAS VIII. American Society of Limnology and Oceanography.

Jaffé, R. and others 2013. Global Charcoal Mobilization from Soils via Dissolution and Riverine Transport to the Oceans. Science **340:** 345-347.

Jaspers, E., K. Nauhaus, H. Cypionka, and J. Overmann. 2001. Multitude and temporal variability of ecological niches as indicated by the diversity of cultivated bacterioplankton. Fems Microbiology Ecology **36:** 153-164.

Jiao, N. and others 2010. Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. Nat Rev Microbiol **8:** 593-599.

Jiao, N. Z., and Q. Zheng. 2011. The Microbial Carbon Pump: from Genes to Ecosystems. Appl. Environ. Microbiol. **77:** 7439-7444.

Jumars, P. A., D. L. Penry, J. A. Baross, M. J. Perry, and B. W. Frost. 1989. Closing the microbial loop: dissolved carbon pathway to heterotrophic bacteria from incomplete ingestion, digestion and absorption in animals. Deep Sea Research Part A. Oceanographic Research Papers **36:** 483-495.

Keil, R. G., D. B. Montlucon, F. G. Prahl, and J. I. Hedges. 1994. Sorptive Preservation of Labile Organic Matter in Marine Sediments. Nature **370:** 549-552.

Kieber, R. J., X. L. Zhou, and K. Mopper. 1990. Formation of Carbonyl Compounds from UV-induced Photodegradation of Humic Substances in Natural Waters - Fate of Riverine Carbon in the Sea. Limnology and Oceanography **35:** 1503-1515.

Killops, S. D., and V. J. Killops. 1993. An Introduction or Organic Geochemistry. Longman Scientific & Technical.

Kim, T. H., H. Waska, E. Kwon, I. G. N. Suryaputra, and G. Kim. 2012. Production, degradation, and flux of dissolved organic matter in the subterranean estuary of a large tidal flat. Mar. Chem. **142:** 1-10.

Kirchman, D. L., A. I. Dittel, S. E. G. Findlay, and D. Fischer. 2004. Changes in bacterial activity and community structure in response to dissolved organic matter in the Hudson River, New York. Aquatic Microbial Ecology **35:** 243-257.

Koch, B. P., T. Dittmar, M. Witt, and G. Kattner. 2007. Fundamentals of Molecular Formula Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter. Analytical Chemistry **79:** 1758-1763.

Koch, B. P., M. R. Witt, R. Engbrodt, T. Dittmar, and G. Kattner. 2005. Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Geochimica Et Cosmochimica Acta **69:** 3299-3308.

Koivusalo, M., P. Haimi, L. Heikinheimo, R. Kostiainen, and P. Somerharju. 2001. Quantitative determination of phospholipid compositions by ESI-MS: effects of acyl chain length, unsaturation, and lipid concentration on instrument response. J Lipid Res **42:** 663-672.

Kritzberg, E., J. Arrieta, and C. Duarte. 2010. Temperature and phosphorus regulating carbon flux through bacteria in a coastal marine system. Aquatic Microbial Ecology **58:** 141-151.

Kujawinski, E. B. 2002. Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS): characterization of complex environmental mixtures. Environ. Forensics **3:** 207-216.

Landa, M., M. Cottrell, D. Kirchman, S. Blain, and I. Obernosterer. 2013. Changes in bacterial diversity in response to dissolved organic matter supply in a continuous culture experiment. Aquatic Microbial Ecology **69:** 157-168.

Lawrence, C. R., and J. C. Neff. 2009. The contemporary physical and chemical flux of aeolian dust: A synthesis of direct measurements of dust deposition. Chem. Geol. **267:** 46-63.

Lechtenfeld, O. J., G. Kattner, R. Flerus, S. L. Mccallister, P. Schmitt-Kopplin, and B. P. Koch. 2014. Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean. Geochimica Et Cosmochimica Acta **126:** 321-337.

Levitus, S., J. I. Antonov, T. P. Boyer, and C. Stephens. 2000. Warming of the World Ocean. Science **287:** 2225-2229.

Liu, Z., R. L. Sleighter, J. Zhong, and P. G. Hatcher. 2011. The chemical changes of DOM from black waters to coastal marine waters by HPLC combined with ultrahigh resolution mass spectrometry. Estuarine, Coastal and Shelf Science **92:** 205-216.

Lobbes, J. M., H. P. Fitznar, and G. Kattner. 2000. Biogeochemical characteristics of dissolved and particulate organic matter in Russian rivers entering the Arctic Ocean. Geochimica Et Cosmochimica Acta **64:** 2973-2983.

Loh, A. N., and J. E. Bauer. 2000. Distribution, partitioning and fluxes of dissolved and particulate organic C, N and P in the eastern North Pacific and Southern Oceans. Deep Sea Research Part I: Oceanographic Research Papers **47:** 2287-2316.

Loh, A. N., J. E. Bauer, and E. A. Canuel. 2006. Dissolved and particulate organic matter source-age characterization in the upper and lower Chesapeake Bay: A combined isotope and biochemical approach. Limnology and Oceanography **51:** 1421-1431.

Lonborg, C., M. Middelboe, and C. P. D. Brussaard. 2013. Viral lysis of Micromonas pusilla: impacts on dissolved organic matter production and composition. Biogeochemistry **116:** 231-240.

Malmstrom, R. R., R. P. Kiene, M. T. Cottrell, and D. L. Kirchman. 2004. Contribution of SAR11 bacteria to dissolved dimethylsulfoniopropionate and amino acid uptake in the North Atlantic ocean. Appl. Environ. Microbiol. **70:** 4129-4135.

Matsumoto, K. 2007. Radiocarbon-based circulation age of the world oceans. Journal of Geophysical Research: Oceans **112:** C09004.

Mayer, L. M. 1994. Surface area control of organic carbon accumulation in continental shelf sediments. Geochimica Et Cosmochimica Acta **58:** 1271-1284.

Mccallister, S. L., J. E. Bauer, H. W. Ducklow, and E. A. Canuel. 2006. Sources of estuarine dissolved and particulate organic matter: A multi-tracer approach. Organic Geochemistry **37:** 454-468.

Mccarthy, M., J. Hedges, and R. Benner. 1996. Major biochemical composition of dissolved high molecular weight organic matter in seawater. Mar. Chem. **55:** 281-297.

Middelboe, M., and N. O. G. Jørgensen. 2006. Viral lysis of bacteria: an important source of dissolved amino acids and cell wall compounds. Journal of the Marine Biological Association of the United Kingdom **86:** 605-612.

Middelburg, J. J. 2011. Chemoautotrophy in the ocean. Geophys. Res. Lett. **38**.

Mladenov, N., M. W. Williams, S. K. Schmidt, and K. Cawley. 2012. Atmospheric deposition as a source of carbon and nutrients to an alpine catchment of the Colorado Rocky Mountains. Biogeosciences **9:** 3337-3355.

Mopper, K., A. Stubbins, J. D. Ritchie, H. M. Bialk, and P. G. Hatcher. 2007. Advanced instrumental approaches for characterization of marine dissolved organic matter: Extraction techniques, mass spectrometry, and nuclear magnetic resonance spectroscopy. Chemical Reviews **107:** 419-442.

Mopper, K., X. Zhou, R. J. Kieber, D. J. Kieber, R. J. Sikorski, and R. D. Jones. 1991. Photochemical degradation of dissolved organic carbon and its impact on the oceanic carbon cycle. Nature **353:** 60-62.

Moran, M. A., and R. G. Zepp. 1997. Role of photoreactions in the formation of biologically labile compounds from dissolved organic matter. Limnology and Oceanography **42:** 1307-1316.

Mullis, K. B., and F. A. Faloona. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol **155:** 335-350.

Murray, A. E., J. T. Hollibaugh, and C. Orrego. 1996. Phylogenetic compositions of bacterioplankton from two California estuaries compared by denaturing gradient gel electrophoresis of 16S rDNA fragments. Appl Environ Microbiol **62:** 2676-2680.

Muyzer, G., E. C. Dewaal, and A. G. Uitterlinden. 1993. Profiling of Complex Microbial-Populations by Denaturing Gradient Gel-Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes-Coding for 16s Ribosomal-Rna. Appl. Environ. Microbiol. **59:** 695-700.

Myklestad, S. M. 1995. Release of extracellular products by phytoplankton with special emphasis on polysaccharides. Science of the Total Environment **165:** 155-164.

Naganuma, T., S. Konishi, T. Inoue, T. Nakane, and S. Sukizaki. 1996. Photodegradation or photoalteration? Microbial assay of the effect of UV-B on dissolved organic matter. Marine Ecology Progress Series **135:** 209-310.

Nagata, T. 2008. Organic Matter–Bacteria Interactions in Seawater, p. 207-241. Microbial Ecology of the Oceans. John Wiley & Sons, Inc.

Nagata, T., and D. L. Kirchman. 1992. Release of dissolved organic matter by heterotrophic protozoa: implications for microbial food webs. Arch. Hydrobiol. Beih. Ergeb. Limnol. **35:** 99-109.

Nübel, U. and others 1996. Sequence heterogeneities of genes encoding 16S rRNAs in Paenibacillus polymyxa detected by temperature gradient gel electrophoresis. J Bacteriol **178:** 5636-5643.

Ogawa, H., and E. Tanoue. 2003. Dissolved organic matter in oceanic waters. Journal of Oceanography **59:** 129-147.

Onumpai, C., S. Kolida, E. Bonnin, and R. A. Rastall. 2011. Microbial utilization and selectivity of pectin fractions with various structures. Appl Environ Microbiol **77:** 5747-5754.

Opsahl, S., and R. Benner. 1997. Distribution and cycling of terrigenous dissolved organic matter in the ocean. Nature **386:** 480-482.

Pinhassi, J. and others 1999. Coupling between bacterioplankton species composition, population dynamics, and organic matter degradation. Aquatic Microbial Ecology **17:** 13-26.

Piontek, J., M. Lunau, N. Handel, C. Borchard, M. Wurst, and A. Engel. 2010. Acidification increases microbial polysaccharide degradation in the ocean. Biogeosciences **7:** 1615-1624.

Pomeroy, L. R. 1974. The Ocean's Food Web, A Changing Paradigm. Bioscience **24:** 499-504.

Pütter, A. 1907. Der Stoffhaushalt des Meeres. Zeitschrift für Allgemeine Physiologie **7:** 321-368.

Redfield, A. C. 1934. On the proportions of organic derivations in sea water and their relation to the composition of plankton, p. 177-192. *In* R. J. Daniel [ed.], James Johnstone Memorial Volume. University Press.

Riebesell, U., J. P. Gattuso, T. F. Thingstad, and J. J. Middelburg. 2013. Preface "Arctic ocean acidification: pelagic ecosystem and biogeochemical responses during a mesocosm study". Biogeosciences **10:** 5619-5626.

Ronaghi, M., M. Uhlén, and P. Nyrén. 1998. A Sequencing Method Based on Real-Time Pyrophosphate. Science **281:** 363-365.

Sambrotto, R. N. and others 1993. Elevated consumption of carbon relative to nitrogen in the surface ocean. Nature **363:** 248-250.

Šantl-Temkiv, T. and others 2013. Hailstones: A Window into the Microbial and Chemical Inventory of a Storm Cloud. PLoS ONE **8:** e53550.

Satterberg, J., T. S. Arnarson, E. J. Lessard, and R. G. Keil. 2003. Sorption of organic matter from four phytoplankton species to montmorillonite, chlorite and kaolinite in seawater. Mar. Chem. **81:** 11-18.

Schlesinger, W. H., and J. M. Melack. 2011. Transport of organic carbon in the world's rivers. Tellus A **33**.

Shimizu, S., T. Yokoyama, T. Akiyama, and Y. Matsumoto. 2012. Reactivity of lignin with different composition of aromatic syringyl/guaiacyl structures and erythro/threo side chain structures in beta-O-4 type during alkaline

delignification: as a basis for the different degradability of hardwood and softwood lignin. J Agric Food Chem **60:** 6471-6476.

Sigman, D. M., and G. H. Haug. 2003. 6.18 - The Biological Pump in the Past, p. 491-528. *In* H. D. Holland and K. K. Turekian [eds.], Treatise on Geochemistry. Pergamon.

Siqueira, J. F., Jr., A. F. Fouad, and I. N. Rocas. 2012. Pyrosequencing as a tool for better understanding of human microbiomes. J Oral Microbiol **4:** 23.

Skoog, A., and R. Benner. 1997. Aldoses in various size fractions of marine organic matter: Implications for carbon cycling. American Society of Limnology and Oceanography.

Smith, D. C., G. F. Steward, R. A. Long, and F. Azam. 1995. Bacterial Mediation of Carbon Fluxes During a Diatom Bloom in a Mesocosm. Deep-Sea Research Part II-Topical Studies in Oceanography **42:** 75-97.

Stenson, A. C., W. M. Landing, A. G. Marshall, and W. T. Cooper. 2002. Ionization and fragmentation of humic substances in electrospray ionization Fourier transform-ion cyclotron resonance mass spectrometry. Analytical Chemistry **74:** 4397-4409.

Stubbins, A. and others 2010. Illuminated darkness: Molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry. Limnology and Oceanography **55:** 1467-1477.

Suttle, C. A. 1994. The significance of viruses to mortality in aquatic microbial communities. Microbial Ecology **28:** 237-243.

Taniguchi, M., W. C. Burnett, J. E. Cable, and J. V. Turner. 2002. Investigation of submarine groundwater discharge. Hydrological Processes **16:** 2115-2129.

Teeling, H. and others 2012. Substrate-Controlled Succession of Marine Bacterioplankton Populations Induced by a Phytoplankton Bloom. Science **336:** 608-611.

Teira, E., H. Van Aken, C. Veth, and G. J. Herndl. 2006. Archaeal uptake of enantiomeric amino acids in the meso- and bathypelagic waters of the North Atlantic. Limnology and Oceanography **51:** 60-69.

Thomas, D. N., and R. J. Lara. 1995. Photodegradation of Algal Derived Dissolved Organic Carbon. Marine Ecology Progress Series **116:** 309-310.

Tranvik, L., and S. Kokalj. 1998. Decreased biodegradability of algal DOC due to interactive effects of UV radiation and humic matter. Aquatic Microbial Ecology **14:** 301-307.

Tranvik, L. J. 1992. Allochthonous dissolved organic matter as an energy source for pelagic bacteria and the concept of the microbial loop. Hydrobiologia **229:** 107-114.

Verdugo, P., A. L. Alldredge, F. Azam, D. L. Kirchman, U. Passow, and P. H. Santschi. 2004. The oceanic gel phase: a bridge in the DOM–POM continuum. Mar. Chem. **92:** 67-85.

Vetter, T. A., E. M. Perdue, E. Ingall, J. F. Koprivnjak, and P. H. Pfromm. 2007. Combining reverse osmosis and electrodialysis for more complete recovery of dissolved organic matter from seawater. Separation and Purification Technology **56:** 383-387.

Volk, T., and M. I. Hoffert. 1985. Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO₂ changes, p. 99-110. The Carbon Cycle and Atmospheric CO₂: Natural Variations Archean to Present. Geophys. Monogr. Ser. AGU.

Wakeham, S. G., T. K. Pease, and R. Benner. 2003. Hydroxy fatty acids in marine dissolved organic matter as indicators of bacterial membrane material. Organic Geochemistry **34:** 857-868.

Weinbauer, M. G., and P. Peduzzi. 1995. Effect of virus-rich high molecular weight concentrates of seawater on the dynamics of dissolved amino acids and carbohydrates. Marine Ecology Progress Series **127:** 245-253.

Wetz, M. S., and P. A. Wheeler. 2007. Release of dissolved organic matter by coastal diatoms. Limnology and Oceanography **52:** 798-807.

Whitman, W. B., D. C. Coleman, and W. J. Wiebe. 1998. Prokaryotes: The unseen majority. Proceedings of the National Academy of Sciences **95:** 6578-6583.

Wilhelm, S. W., and C. A. Suttle. 1999. Viruses and Nutrient Cycles in the Sea - Viruses play critical roles in the structure and function of aquatic food webs. Bioscience **49:** 781-788.

Williams, P. M., and E. R. M. Druffel. 1987. Radiocarbon in dissolved organic matter in the central North Pacific Ocean. Nature **330:** 246-248.

Wilmes, P., and P. L. Bond. 2009. Microbial community proteomics: elucidating the catalysts and metabolic mechanisms that drive the Earth's biogeochemical cycles. Curr. Opin. Microbiol. **12:** 310-317.

Witt, M., J. Fuchser, and B. P. Koch. 2009. Fragmentation Studies of Fulvic Acids Using Collision Induced Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Analytical Chemistry **81:** 2688-2694.

Woods, G. C., M. J. Simpson, P. J. Koerner, A. Napoli, and A. J. Simpson. 2011. HILIC-NMR: Toward the Identification of Individual Molecular Components in Dissolved Organic Matter. Environmental Science & Technology **45:** 3880-3886.

Zhan, A. and others 2013. High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. Methods in Ecology and Evolution **4:** 558-565.

Zinger, L. and others 2011. Global Patterns of Bacterial Beta-Diversity in Seafloor and Seawater Ecosystems. PLoS ONE **6**.

Curriculum Vitae

Name	Helena Osterholz
Date of birth	Dec. 26, 1984
Place of birth	Paderborn, Germany

Education

2004-2007	B.Sc. Environmental Sciences, University of Bielefeld
2007-2010	M.Sc. Marine Environmental Sciences, Carl von Ossietzky University Oldenburg
2010-2014	Ph.D. student Research Group for Marine Geochemistry, ICBM Oldenburg/MPI Bremen

Publications

Osterholz, H., H. Simon, M. Beck, J. März, S. Rackebrandt, H.-J. Brumsack, U. Feudel, M. Simon (2014) Impact of diatom growth on trace metal dynamics (Mn, Mo, V, U). *Journal of Sea Research* 87: 35-45

Osterholz, H., T. Dittmar, J. Niggemann (2014) Molecular evidence for rapid dissolved organic matter turnover in Arctic fjords. *Marine Chemistry* 160: 1-10

Conference Contributions

Osterholz H., J. Niggemann, C. Feenders, B. Blasius, M. Simon, T. Dittmar. Fast transformation of freshly produced compounds into refractory dissolved organic matter. YOUMARES 4. German Young Marine Scientist Meeting, Oldenburg, Germany, Sept. 11-13, 2013 (Talk).

Osterholz H., S. Müllenmeister, C. Kramer, M. Simon, J. Niggemann, T. Dittmar. Fast transformation of freshly produced compounds into refractory dissolved organic matter. Association for the Sciences of Limnology and Oceanography, New Orleans, USA, Feb. 17-22, 2013 (Talk).

Osterholz, H., B. Engelen, B. Fuchs, T. Dittmar, J. Niggemann. Molecular evidence for rapid dissolved organic matter turnover in Arctic fjords. 3rd Meeting of SCOR WG 134 – The Microbial Carbon Pump in the Ocean, Aug. 27-28, 2012, Delmenhorst, Germany (Poster).

Osterholz, H., B. Engelen, J. Niggemann, T. Dittmar. Linking Dissolved Molecules and Microorganisms – A Case Study in Arctic Fjords. YOUMARES 2. German Young Marine Scientist Meeting, Bremerhaven, Germany, Sept. 7-9, 2011. (Talk).

Research Cruises and Fieldwork

R/V Polarstern	ANTXXVIII-II, Cape Town – Neumayer 3 – Cape Town (Dec. 3 2011-Jan. 5, 2012)
R/V Heincke	HE361, German Bight and North Atlantic (July 12-July 29, 2011)
MS FARM	Svalbard fjords, Norway (Aug. 31-Sept. 15, 2010)

Acknowledgements

First of all, I would like to thank Prof. Dr. Thorsten Dittmar for the opportunity to take on the challenge of this thesis in his working group and for the kind supervision. Prof. Dr. Meinhard Simon I thank for reviewing this thesis, but more importantly for the guidance over many years. I would also like to thank Prof. Dr. Hans Brumsack for agreeing to be part of my doctorate committee.

A special thank you I would like to express towards Dr. Jutta Niggemann for supervising me, including an exciting cruise, many discussions, and patiently reading drafts of posters and manuscripts again and again.

I am indebted to Dr. Gabriel Singer for letting me in on some secrets of multivariate statistics and Dr. Jeffrey Hawkes is thanked for improving opening and closing sections of this thesis.

The atmosphere among the “geochems” has always been great. Especially the DOM/R-support group and the Aquarium office mates ought to be acknowledged. Thank you, Maren Seibt, Marcus Manecki, Michael Seidel, Hannelore Waska, Maren Stumm and Pamela Rossel. Susanne Wendeling simplified all administrative matters a great deal.

Katrin Klaproth, Matthias Friebe and Ina Ulber I thank for technical support. Not to be overlooked are the open doors everywhere at the ICBM. Many people, among them Heike Simon, Dr. Bert Engelen, Dr. Bernhard Schnetger, Carola Lehners, Dr. Helge Ansgar Giebel, Matthias Wolterink, Rolf Weinert, Birgit Kürzel, Andrea Schlingloff, Michael Pilzen, Dr. Christoph Feenders and many others trustfully let me use their labs, helped with analyses or discussions and ideas.

Furthermore, I thank the crews and scientific parties of the 2012 Svalbard expedition, Heincke HE361 and Polarstern cruise ANTXXVIII/2 who helped me obtain the data for this thesis, but also contributed to many fun memories made during my time as a PhD student.

Yvonne Lipsewers, Judith Lucas, Moni Sahlberg, Tim Engelhardt – thank you for the distraction and good times during the long hours in the kitchen.

Sascha, thank you for your understanding and support, for being there.

My family, I am indebted to you for your everlasting support that enables me to follow my own way. Thank you!

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Zusätzlich erkläre ich, dass diese Dissertation weder in ihrer Gesamtheit noch in Teilen einer anderen wissenschaftlichen Hochschule zur Begutachtung in einem Promotionsverfahren vorliegt oder vorgelegen hat.

Oldenburg, im Juni 2014 Helena Osterholz