Cycling of bioavailable carboxyl-rich alicyclic molecules and carbohydrates in Baffin Bay

Student: Kayla McKee

Supervisor: Dr. Brett Walker

In partial fulfillment of the requirements for the M.Sc. degree in Earth Sciences

Department of Earth and Environmental Sciences Faculty of Science

University of Ottawa

uOttawa

Table of Contents

Statement of Authorship and Collaborator Contributions

This thesis comprises a collaboration with co-authors Dr. Hussain Abdulla (H.A.), Lauren O'Reilly (L.O.), and supervisor Dr. Brett Walker (B.D.W.). All co-authors declare no conflict of interest through their collaboration in this thesis. We declare no ethics approval was required. K.M. wrote the paper with input from B.D.W. Comments were provided by H.A. and L.O. ^{1}H-NMR analysis, data reduction and integrations were performed by K.M. Data calculations and figure production was performed by K.M. Total DOC concentration measurements were performed by L.O. We would also like to thank Dr. Brett Walker for conceptualizing the project, providing the necessary facilities, and overseeing the project. We would like to thank Dr. Hussain Abdulla for baseline correction training, data integration training, and offering his expertise.

Thesis Abstract

At ~662 gigatonnes of carbon (GtC), marine dissolved organic matter (DOM) is the largest reduced pool of actively cycling carbon and nitrogen in the oceans[1]. Operationally defined as smaller than 0.1µm in size, this carbon reservoir comprises all non-living organic matter smaller than a bacterial cell and comprises organic colloids and molecules spanning as a continuum of sizes ranging from marine viruses and large macromolecules (e.g. DNA, enzymes) to small organic molecules (e.g. polymers and monomers)[2]. With deep apparent ^{14}C-ages ranging between 4900-6400 ybp [3,4], marine DOM is anomalously old given timescales of global ocean ventilation (1000-1500 years). The great age of DOM has remained one of the most elusive lines of scientific inquiry in Chemical Oceanography for decades. The size and molecular composition of DOM has been shown to be a key variable in determining its biological reactivity (e.g. cycling rate) and long-term persistence in the deep ocean[5,6].

Despite the importance of DOM in the marine carbon and nitrogen cycles, we lack a detailed understanding of the molecular composition of DOM. Due to the high concentration of salts in seawater relative to DOM, it is difficult to analyze the molecular composition of seawater with conventional chemical- or size- fractionation methods without introducing bias (i.e. isolating only hydrophobic and/or high molecular weight DOM). In fact, it is commonly reported that >80% of DOM remains uncharacterized at the molecular level (e.g. not readily identifiable as an individual known biomolecule)[5]. Nuclear magnetic resonance (NMR) spectroscopy has been used as a tool for several decades to describe the composition of marine DOM isolates[7]. For example, ^{13}C-NMR of major high molecular weight DOM functional groups at the molecular-level demonstrated that DOM is largely made up of reactive polysaccharides with low aromaticity compared to terrestrial DOM[8]. To date, all marine DOM NMR measurements have been made on size-fractionated DOM or chemically-fractionated (e.g. solid phase extracted) DOM isolates. In this thesis, I report the first Proton (^{1}H) NMR composition of total seawater DOM from seawater samples collected from 10 stations in Baffin Bay aboard the CCGS *Amundsen* (2019). Samples were measured using ^{1}H-NMR at uOttawa following a novel water suppression method established by Lam and Simpson[9].

The use of this method has allowed for the first molecular composition assessment of total seawater DOM to be measured (e.g. without any chemical or size fractionation). I report the %

relative abundance of individual biomarkers and determine molar concentrations of two compound classes of interest. These results are shown in Ocean Data View section plots, and are listed within appendix tables, to provide a comprehensive depiction of the changing concentrations of dissolved organic carbon (DOC), total carbohydrates (TCHO), and carboxyl-rich alicyclic molecules (CRAM). In this thesis, I explore changes in the abundance of these unique DOM compound classes and discuss how the composition of DOM directly determines its bioavailability and thus cycling in Baffin Bay [5]. The core objective of my thesis was to measure DOM concentrations for TCHO and CRAM, as well as to calculate the production and removal of these key DOM compounds in Baffin Bay due to either physical and/or biological processes. We found that the concentration of both TCHO and CRAM decreased with depth throughout Baffin Bay. This is consistent with previous work suggesting the rapid cycling of carbohydrates, however it contradicts the current paradigm of CRAM cycling. Our results indicate between 21-43% of CRAM produced in the surface is subsequently removed at depth. Rapid cycling of a surface CRAM population suggests that not all CRAM can be considered recalcitrant DOM

We live in a time of unprecedented global change. The Arctic Ocean is warming at a rate at least four times faster than the global average[10]. The impact of a rapidly warming, freshening and increasingly acidified Arctic Ocean on the biogeochemistry of DOM remains unknown. It is imperative that more DOM research be conducted as early as possible in order to better understand these impacts and inform future research directions. The distribution and cycling of CRAM in Baffin Bay provide novel and fundamental knowledge of DOM cycling in a key Arctic region, but could also potentially occur throughout the global ocean. Such data will no doubt be of use in informing future iterations of Earth System Climate models seeking to forecast how the marine carbon cycle will respond to global change.

Thesis Acknowledgements

The completion of this thesis could not have been possible without the assistance and support of my mentors, colleagues, collaborators, and friends. Their contributions are deeply appreciated and greatly acknowledged.

Firstly, I would like to thank my supervisor Dr. Brett Walker, for his continuous support and guidance through my undergraduate and graduate degree. His patience, motivation, and genuine enthusiasm have made this process both enjoyable and memorable. Dr. Walker's expertise and excellent teaching skills have provided me with important skills in scientific research that I will be able to apply throughout the rest of my career. Above all, I would like to thank Dr. Walker for his kindness, friendship, and encouraging nature, and I am extremely grateful for all the opportunities he has provided me with.

I would also like to thank my co-supervisor, Dr. Hussain Abdulla, for providing his own expertise, guidance, and support throughout my studies. The completion of this research project would not have been possible without his help throughout the years.

I gratefully acknowledge Amundsen Science staff, Chief Scientist Alexandre Forest, Anissa Merzouk and the crew of the CCGS *Amundsen* for their support and the opportunity to participate in the 2019 research expedition. I am also grateful for the NMR facility staff Dr. Glenn Facey, Dr. Peter J. Pallister, and Dr. Vincent Morin for providing me with training for the ^{1}H-NMR instrument and their ongoing support while running my samples. I acknowledge support from the ArcticNet Network of Centers of Excellence of Canada and Amundsen Science as a Major Scientific Initiative Centre. I would like to thank Sara Zeidan for aiding in sample collection. I would also like to thank Claire Normandeau and Qiang Shi of Dalhousie University for performing high temperature combustion DOC measurements.

I am also extremely grateful to my family and friends who have always supported and motivated me to complete my research. Without their unconditional love and care, none of this would have been possible.

List of Figures

List of Tables

Glossary

CHO: Carbohydrates

CRAM: Carboxyl-Rich Alicyclic Molecules

CP-MAS: Cross-Polarization Magic-Angle-Spinning

^{13}C-NMR: Carbon 13 Nuclear Magnetic Resonance

DOC: Dissolved Organic Carbon

DOM: Dissolved Organic Matter

DCNS: Dissolved Combined Neutral Sugars

HDPE: High Density Polyethylene

HMW: High Molecular Weight

^{1}H-NMR: Proton Nuclear Magnetic Resonance

HPS: Heteropolysaccharides

HSQC: Heteronuclear Single Quantum Coherence

LDOM: Labile Dissolved Organic Matter

LMW: Low Molecular Weight

MBTH: 3-Methyl-2-Benzothiazoline Hydrazone

MCP: Microbial Carbon Pump

NMR: Nuclear Magnetic Resonance

PCA: Principal Component Analysis

RDOM: Recalcitrant Dissolved Organic Matter

SPE: Solid Phase Extracted

TCHO: Total Carbohydrates

TOC: Total Organic Carbon

TMS: Tetramethylsaline

UF: Ultrafiltered

1. Background and Review of Relevant Literature

1.1 The Biogeochemistry of DOM in Arctic Systems

The Arctic Ocean is a globally significant ocean region that links the Atlantic and Pacific Oceans. It is also an area of large dissolved organic matter (DOM) production and export. The net result of mixing Atlantic and/or Pacific waters, riverine input, *in situ* production, degradation, and sea ice melt all control the distribution and concentration of dissolved organic carbon (DOC) in the Arctic.

Both Pacific and Atlantic surface currents transport significant amounts of nutrients to Arctic shelves and marginal seas. This high nutrient supply and stabilized water column result in distinctly high rates of primary production[11–13]. The recent average estimate for net primary productivity in the pan Arctic was 101 gC m^{-2} $year^{-1}$ or 440 TgC per year[14]. Approximately 10% of this productivity accumulates as DOC[15]. The central Arctic Ocean has low primary productivity compared to the shelf seas[16]. The central Arctic Ocean was found to produce 0.5 mol m^{-2} $year^{-1}$ which will result in an increase of 10 µmol C L^{-1} in the top 50m [17]. This rate of DOC production cannot be sustained over the residence time of surface water (5-10 years) without DOC removal, otherwise DOC would build up to unrealistically high concentrations. Thus, there must be extensive recycling and removal of DOC to balance seasonal production in the Arctic.

The Arctic Ocean receives approximately 10% of the global river discharge[18] with the largest contributors being the large Siberian Rivers and Mackenzie River. This runoff has the highest concentration of DOC in Arctic Ocean source water and thus contributes a large amount of terrigenous DOM to the surface layer of the Arctic Ocean[19–21]. This large terrestrial component of DOM is unique to the Arctic Ocean. Given the high seasonal variability however, it remains difficult to meaningfully quantify the annual input of DOC into the Arctic Ocean. Based on the

conservative mixing behavior as well as long term decomposition studies it is thought that the DOC in major Arctic rivers is largely refractory[22] however this might not always be the case[23–26]. For example, deep water Arctic DOC concentrations are much lower than the riverine DOC concentrations entering the ocean suggesting only very small amounts of terrigenous DOM can be exported to the deep Arctic[27].

DOC accounts for approximately 5% of total carbon fluxes in the Arctic Ocean and is negatively correlated to water mass salinity. The upper (50-600m) Atlantic Ocean inflow water has an average DOC concentration of 58 ± 5 µmol C L^{-1} [28]but can vary from 52-108 µmol C L^{-1} [17,29–32]. While the average is comparable to the Pacific Ocean, it is an order of magnitude less than the DOC input from the major river systems that flow into the Arctic Ocean[25,32–34]. Surface DOC concentrations, measured by Shen and co-workers, in Baffin Bay ranges from 66-97 µmol C L^{-1} (77 ± 10 µmol C L^{-1}) [35,] however to the best of our knowledge, no water column depth profiles of DOC concentrations have been reported. Using a dissolved D/L amino acid proxy, Shen and co-workers found significant export of bioavailable DOM within the Canadian Arctic Archipelago and in Baffin Bay[35], largely driven by light availability and primary productivity. The long transit time of water masses through the Arctic Archipelago provides ample time for microbial degradation which could contribute to the low bioavailability of DOM in some locations in Baffin Bay. In contrast, the Greenland coast is host to warmer water and less sea ice during the summer months and as a result Shen et al. found elevated amino acid yields indicating the presences of DOM with high bioavailability[35]. Shen hypothesized that the Greenland coastal water may have nutrient input from glacial meltwater. In the pan-Arctic, between 1-5% of DOM is estimated to be labile (cycling on timescales of hours-weeks), 26-30% is semi labile (months-years), and

approximately 70% is refectory (decades)[35]. However, the linkages between DOM composition and cycling rates within the Arctic remains largely unknown.

1.2 Nuclear Magnetic Resonance and the Emerging Picture of DOM Composition

DOM is a globally significant carbon (662 GtC) reservoir and plays an important role in the marine carbon cycle. Marine DOM is operationally defined as material smaller than 0.1µm and is comprised of millions of different molecules, many of which reside in the oceans for thousands of years[36]. Chemical analysis of DOM can provide insight into the role of DOM in biogeochemical cycles and climate attenuation, interaction with microorganisms, its lifespan, and its molecular composition[36]. Interest in understanding the biological reactivity and cycling of DOM has gained renewed attention in the field of Chemical Oceanography following technological advances in our ability to measure its composition resulting in a new understanding of the microbial carbon pump[37] and the production of recalcitrant DOM by heterotrophic marine bacteria.

Nuclear magnetic resonance (NMR) spectroscopy is a versatile and nondestructive analytical tool. It is currently the only technique that allows the user to determine the structure of molecules in solution[38]. NMR spectroscopy subjects the sample to a magnetic field which excites positively charged nuclei (e.g. ^{1}H, ^{13}C) and causes their randomly distributed spin orientations to align in parallel with the external magnetic field. These nuclei are subjected to electromagnetic radiation to induce “spin flips” from low to high energy states and produce resonance. Small differences in the effective magnetic fields for electrons orbiting these nuclei result in a series of distinct, yet absolute, NMR signals (chemical shifts, (d_H) ppm), allowing NMR to effectively map the hydrogen-carbon framework of a molecule[39]. In this thesis I use proton (^{1}H) NMR as a

quantitative tool to identify changes in marine DOM functional group composition (e.g. compound classes) present in seawater.

NMR has played a key role in shaping our understanding of DOM composition and cycling. Early studies used tangential flow ultrafiltration (UF) to isolate large quantities of DOM greater than 1nm in size, commonly known as high molecular weight (HMW) DOM. When coupled with diafiltration, this allowed for both more detailed isotopic and chemical characterization of salt-free DOM. Early ^{13}C-NMR measurements suggested approximately 50% of surface HMW DOM was comprised of polysaccharides (carbohydrates). However, HMW DOM also contains significant quantities of aliphatic molecules and carboxylic acids, in particular for deep HMW DOM[8,40]. Surface carbohydrates were then hypothesized to comprise a carbon-rich substrate that could support marine heterotrophy and could build up in subsurface waters as it was subsequently degraded over time. ^{1}H-NMR measurements added to the DOM composition and cycling picture by showing that neutral sugars, acetate, and lipids followed a similar depth trend, suggesting surface autotrophic production and rapid heterotrophic removal of complex macromolecular heteropolysaccharides (HPS; eg. N-acetyl amino polysaccharides and acyl oligosaccharides)[41,42]. Subsequent mass-spectrometry studies revealed major HPS components are comprised of a suite of methylated hexoses–, pentoses–, 6-deoxysugars– (e.g. O-Methylrhamnose or O-Methylfucose), heptoses–, 3,6-dideoxysugars, and 1,6 anhydrosugars[43]. These findings are supported by independent methods used for determining dissolved combined neutral sugar (DCNS) and total carbohydrate (TCHO) concentrations from seawater. Pacific and Atlantic DCNS and TCHO distributions have revealed that DOM sugars cycle on seasonal timescales as bioavailable DOM[44]. We now have a robust understanding of bioavailable or labile DOM (LDOM) cycling through several biomolecular proxies (e.g. the concentrations of individual sugars, fatty acids, amino acids,

etc.), however little is known about the composition and cycling of refractory DOM (RDOM), which may persist in the ocean for centennial to millennia.

Early Cross-Polarization Magic-Angle-Spinning (CP-MAS) ^{13}C-NMR measurements showed seawater DOM was high in aliphatic- and carboxyl-carbon assembled in a highly branched and interlinked (possibly cyclic) structure[45]. Taken together with high ^{14}C-ages of deep DOM[46], it was hypothesized that such "humic-like" compounds which accumulate in seawater, represented biologically refractory DOM that remains in the ocean for several mixing cycles[40,47]. Through the use of both ultrahigh resolution mass spectroscopy and novel 2D Heteronuclear Single Quantum Coherence (HSQC; ^{1}H, ^{13}C) NMR experiments, Hertkorn and co-workers recently defined this material as "carboxyl-rich-alicyclic molecules", or CRAM. They found that it comprises ~8% of HMW DOM[48].

Dissolved black carbon, humic-like fluorescent DOM, carotenoids, and deaminated peptides have been identified as representative RDOM [22–25]. However, the formation of CRAM via the microbial carbon pump (MCP) is most commonly referenced as an example of RDOM in the literature[48–50]. Recent NMR measurements of solid phase extracted (SPE) DOM by Hertkorn and co-workers suggest that CRAM comprises between 51-56% of SPE DOM which makes up ~40% of the total DOM pool[51]. More recently, isotopic and NMR studies have implemented a targeted combination of UF and SPE to isolate both HMW DOM (as LDOM) and low molecular weight (LMW) SPE DOM (as RDOM) fractions from two oligotrophic subtropical stations in the Atlantic and Pacific oceans[52–54]. These isotopic and NMR studies demonstrated little difference between LMW SPE-DOM isolated between the ocean basins – strongly supporting the hypothesis that LMW SPE-DOM is representative of RDOM, rich in CRAM[29–31]. A recent long-term microbial incubation study combining ultra-high resolution mass spectrometry and NMR analysis

of SPE DOM from the south China sea found no change in DOM composition over 180 day incubations at all depths[50]. Taken together, these studies support the hypothesis that CRAM-rich SPE-DOM is a representative proxy for RDOM. However, to date, our ability to quantify the RDOM cycle has been impeded by several key factors.

First, all ultra-high resolution mass spectrometry and NMR analyses have relied heavily on chemical- and/or size-fractionation to isolate sufficient salt-free DOM for detailed analysis, however this can only obtain between 40-60%, at most, of the DOM present in the sample. To the best of our knowledge, no %CRAM estimates (as a percent of dissolved organic carbon; DOC) within the total DOM pool have been reported. Second, RDOM is remineralized to CO_2 through photochemical oxidation and/or microbial heterotrophy on timescales of hundreds of years[37]. This complicates the direct observations of RDOM biogeochemical cycling. Third, to the best of our knowledge, existing NMR studies are limited to reporting of qualitative CRAM estimates as "relative percent abundance". To truly understand CRAM cycling and biogeochemistry, quantitative estimates of changes in environmental CRAM "standing stock" concentrations (i.e. as μmol C L^{-1}) are essential for furthering our understanding RDOM biogeochemistry.

Our study aims to tackle the aforementioned short comings of DOM analysis with respect to CRAM. In a standard NMR probe, a ~600 μL seawater DOM sample contains between 500-600 ng C[9]. Therefore, it is necessary to use a water suppression method, such as that set forth by Lam and Simpson[9] to measure the low abundance of organic matter in a sample with a comparatively high abundance of water and salt. By implementing an adapted version of Lam and Simpson's method we were able to measure total seawater DOM composition without the need for chemical- or size-fractionation. Because ^{1}H-NMR is quantitative and the H/C ratio of CRAM is

well parameterized, this allows us to calculate the concentration of CRAM rather than simply the relative percent abundance of the compound in seawater.

Cycling of bioavailable carboxyl-rich alicyclic molecules and carbohydrates in Baffin Bay

Kayla McKee[1], Hussain Abdulla[2], Lauren O'Reilly[1], Brett D. Walker[1*]

[1]Department of Earth and Environmental Sciences, University of Ottawa, Ottawa, ON, Canada

[2]Department of Physical and Environmental Science, Texas A&M University-Corpus Christi, Corpus Christi, TX, United States

[*]Corresponding Author: brett.walker@uottawa.ca

**Manuscript in review at *Nature Communications*

2. Statement of Key Chapter Findings and Relevance

Marine dissolved organic matter (DOM) is an actively cycling, and globally significant carbon reservoir (662 GtC) comprising millions of chemically distinct compounds[1,36]. Despite the importance of DOM in the marine carbon cycle we still lack a detailed understanding of its chemical structure and cycling in the ocean. Of particular importance, is improving estimates of DOM cycling within rapidly warming, polar environments. Previous DOM studies have shown rapid surface cycling of carbohydrates, which is broadly considered biologically labile DOM (LDOM)[5]. Conversely, carboxyl-rich alicyclic molecules (CRAM) are thought to persist at depth as biologically recalcitrant DOM (RDOM)[48]. Previous studies using ultrafiltration and/or solid-phase extraction to isolate salt-free DOM can only recover a small fraction of DOM (<20-40%)[52,53]. These methods induce chemical (e.g. hydrophobicity), size (e.g. >1nm) and isotopic (e.g. ^{13}C, ^{14}C, ^{15}N) bias – making inferences to the cycling of the total DOM reservoir problematic[6,55]. In this study, we use a novel total seawater DOM proton (^{1}H) nuclear magnetic resonance (NMR) spectroscopy method to quantify the concentrations of total carbohydrates and CRAM throughout Baffin Bay in the Canadian Arctic. Our results show that both carbohydrates and CRAM have high concentrations in the surface ocean and low concentrations at depth. This is consistent with previous work suggesting the rapid cycling of carbohydrates, however it contradicts the current paradigm of CRAM cycling. Our results indicate between 21-43% of CRAM produced in the surface is subsequently removed at depth. Rapid cycling of a surface CRAM population suggests that not all CRAM can be considered RDOM. These results provide new constraints on the biogeochemical cycling of two major DOM components and the long-term persistence of DOM in the deep ocean.

3. Introduction

Understanding the nature and cycling of DOM is of paramount importance for constraining the marine carbon cycle. Given the size and great age (4-6kyr) of the DOM reservoir[46], DOM acts as a capacitor for deep ocean carbon storage, mitigating climate on timescales of centuries to millennia. DOM is likely comprised of millions of different compounds[36] yet the majority of DOM remains uncharacterized at the molecular level[56], or elusive to current direct isolation techniques. This has precluded our ability to quantify DOM molecular transformations and cycling. The Arctic Ocean is warming four times faster than the rest of the planet[57]. Resulting ecosystem level impacts on phytoplankton communities and blooms, together with changes in DOM production, are likely to alter the microbial loop and sequestration of DOM in these regions[58]. Baffin Bay is a known region of significant labile DOM (LDOM) production[35] based on Shen's comprehensive study which quantified the amino acid composition of DOM throughout Baffin Bay in 2018. However, little is known about recalcitrant DOM (RDOM) cycling in this region and the biogeochemical cycling of DOM within Baffin Bay remains poorly constrained.

Detailed chemical analysis, and in particular NMR measurements, have shaped our understanding of DOM composition and cycling. Early studies using tangential-flow ultrafiltration (UF) for the first isolation of high molecular weight (HMW; >1nm) DOM allowed for both isotopic and chemical characterization of DOM. Through ^{13}C-NMR measurements it was estimated that ~50% of surface ultrafiltered high molecular weight DOM (HMW DOM; >1000 Da) was carbohydrates whereas both surface and deep HMW DOM were composed of primarily aliphatic molecules and carboxylic acids[8,40]. It was inferred that surface carbohydrates represented a major carbon rich substrate for biological respiration in the upper water column which could then be advected and slowly degraded to subsurface waters. Subsequent ^{1}H-NMR analyses revealed neutral sugars, acetate, and lipids shared similar depth trends, suggesting the biosynthetic surface production and rapid removal of complex macromolecular heteropolysaccharides (HPS; including acyl oligosaccharides, and N-acetyl amino polysaccharides)[41,42]. This was later confirmed through mass-spectrometry studies identifying methylated hexoses–, pentoses–, 6-deoxysugars–, heptoses–, 3,6-dideoxysugars, and 1,6 anhydrosugars as major components of HMW DOM HPS[43]. This work is supported by direct measurements of total carbohydrates (TCHO) and dissolved combined neutral sugars from the Atlantic and Pacific, suggesting this to be a major labile DOM component, cycling on seasonal timescales[44]. While our understanding of bioavailable or LDOM

cycling through its biomolecular proxies (e.g. amino acids, sugars, fatty acids, etc.) are well constrained, little is known about the composition and cycling of refractory DOM on centennial to millennial timescales.

Early NMR studies revealed the refractory nature of marine "humic" substances within the DOM reservoir and in the deep ocean. Deep ocean HMW DOM has high abundances of carboxyl and aliphatic carbon held in highly branched and interlinked (e.g. cyclic) structures[45]. Based off ^{13}C-NMR, this material is found throughout the water column, but in higher relative total abundance in deep HMW DOM, suggesting a refractory "humic" background DOM component that lasts several oceanic mixing cycles[40,47]. Using a combination of NMR and ultrahigh resolution mass spectroscopy, this refractory material was later coined "carboxyl-rich alicyclic material" (CRAM) and was estimated to comprise ~8% of HMW DOM[48] .

Several molecular compound classes have been identified as key components of the long-lived recalcitrant RDOM pool, including: dissolved black carbon, carotenoids, humic-like fluorescent DOM and deaminated peptides[59–62]. However, CRAM is the compound class most commonly referenced as a suitable proxy for RDOM formed via the microbial carbon pump[48–50]. Recent NMR and isotopic studies implemented a combined ultrafiltration and solid phase extraction approach to isolate HMW and low molecular weight solid phase extracted DOM (LMW SPE DOM; <1000 Da) fractions from two oligotrophic subtropical stations in the Pacific and Atlantic (ALOHA and BATS) suggest the composition of LMW SPE-DOM closely matches that of CRAM and that LMW SPE-DOM composition is invariant with depth and between the two ocean basins[52–54]. Recent long-term microbial incubations combining NMR and ultra-high resolution mass spectrometry of SPE-DOM from the south China sea found invariant DOM compositions over 180 day incubations at all depths[50]. Hertkorn and co-workers more recently estimated that CRAM comprises between 51-56% of solid phase extractable (SPE) DOM[51].

Together, these studies highlight SPE-DOM as a suitable proxy for RDOM, one that is rich in CRAM. However, several key problems impede our ability to quantify RDOM cycle. First, the centennial timescales in which RDOM is remineralized via microbial heterotrophy and/or photochemical oxidation[37] complicate direct observations of its biogeochemical cycling. Second, all NMR and ultra-high resolution mass spectrometer studies have relied on size– and/or chemical–fractionation methods, which can obtain at most 40-60% of DOM. To the best of our knowledge,

no %CRAM estimates (as % total dissolved organic carbon; DOC) within the total DOM pool have been reported. Third, NMR has only been used to estimate relative percent abundance of CRAM, precluding precise carbon budgets or estimates of "standing stock" concentrations (e.g. µmol C kg^{-1}) of key RDOM components.

Here we present total DOM ^{1}H-NMR data from 10 stations in Baffin Bay (Figure 1) using a novel water suppression method[9,63]. We use total ^{1}H-NMR abundance data together with DOC concentrations and known functional group H/C ratios to quantify the molar abundance and examine the cycling of two DOM endmembers: total TCHO and CRAM. We evaluate the utility of relative percent abundance estimates and compound concentrations (in µmolC L^{-1}) as novel representatives of LDOM or RDOM throughout Baffin Bay.

4. Methods

4.1 Sample Collection

Seawater samples were collected aboard the CCGS Amundsen on Leg 2a of the Arctic Net expedition in Baffin Bay between July 5th to 25th, 2019 (Figure 1). The locations of the 10 stations were selected to provide comprehensive data for the entire bay; as well, a subset of stations were selected to follow the Western Greenland Current. At sea, samples collected from depths shallower than 400m were filtered using a custom-built 306 Stainless Steel 70mm filter manifold. The manifold was pre-cleaned with 10% HCl and copious amounts of Milli-Q water (3ppb TOC) and dried prior to use. The manifold was loaded with pre-combusted (540°C, 2h) 70mm glass fiber filters in two layers. A 2.0µm quartz filter (Whatman QMA) was placed on top of a 0.7µm borosilicate filter (Whatman GF/F). Dissolved organic carbon (DOC) and ^{1}H-NMR samples were collected by filtering water from 11L Niskin-type Bullister bottles via acid-cleaned PTFE and silicon tubing into pre-combusted, 10mL pre-scored borosilicate ampules (Wheaton part#176780) for ^{1}H-NMR, and into pre-cleaned (overnight 10% HCl soak) 60 mL high density polyethylene (HDPE) bottles (Fisher Scientific, #03-331-32B) for DOC. HDPE bottles were immediately frozen

and stored at -20°C for future analysis. Ampules for ^{1}H-NMR samples were overfilled with 5x sample volume to reduce potential contamination. After collection, ampules were capped with pre-cleaned polypropylene ¼" column caps until they could be poisoned with 1 drop (~50µL) saturated mercuric chloride ($HgCl_2$, Fisher #M1136-100G), and flame sealed with a small butane torch. To further minimize contamination, a separate dram vial of saturated $HgCl_2$ was used for each station. Poisoned and flame-sealed samples were cooled, homogenized, checked for leaks and stored in the dark at room temperature until further sample processing at the University of Ottawa.

4.2 Dissolved Organic Carbon Concentrations

A total of 98 seawater samples were prepared for DOM analysis at the University of Ottawa. Briefly, samples were rapidly thawed, homogenized and transferred into pre-combusted (540°C/2h) 20mL glass autosampler vials (Fisher Scientific, #03-375-25) and acidified to pH < 2 with 50µL of 12M HCl and capped with pre-cleaned (10% HCl) PTFE lined silicone septa caps. Samples were re-frozen, packaged, and shipped to the CERC.OCEAN laboratory (Dalhousie University) for DOC concentration analysis using a Shimadzu TOC-L Total Organic Carbon (TOC) analyzer. A Five Point Calibration was performed prior to sample analysis to create a TOC calibration curve using reagent grade KHP in Milli-Q water. A Milli-Q Water blank and 100µL of DOC Consensus Reference Material (Batch 20 Hansell Lab) was injected prior to the analysis of our samples and was repeated after every six sample analyses for quality control. Samples were injected at least three times, with a maximum of five injections, until a CV better than ±2% was achieved.

4.3 Proton NMR Sample Preparation

Due to the sensitivity of the water suppression ^{1}H-NMR analysis technique and extremely low concentrations of sea water DOM in our samples (40-80 µmol L^{-1}), rigorous cleaning of NMR

tubes and meticulous sample preparation are required. Also, since NMR tubes will bend if baked in an oven and damage the probe, these tubes cannot be pre-combusted prior to analysis. Instead, our NMR tubes are cleaned with Milli-Q water (TOC = 3ppb) and the sample itself. In order to minimize residual particulate organic matter (POM) contamination that interferes with the water suppression ^{1}H-NMR method, ampules were centrifuged at a relative centrifugal force (RCF) of 1900 x G for 20 min and allowed to come to a rest in the centrifuge, gradually, without braking. This effectively eliminated POM contamination in the samples. After centrifugation, the sample ampule was gently cracked open and sample was gently pipetted from the top using a new baked (540°C/2h) 2mL pipet into NMR tubes (Wilmad Precision 500MHz; part # 665000575). The aspiration was done very carefully so as not to resuspend any residual POM in the sample.

Prior to sample loading, NMR tubes were rinsed five times with ~1mL of 18.2 MΩ (Milli-Q) water (TOC = 3ppb) using a pre-baked (540°C, 2h) 2mL Pasteur pipet. For each rinse, ~1mL of Milli-Q water was added, the NMR tube capped and shaken vigorously to ensure the tube was fully rinsed, then decanted into a waste container. The Milli-Q water used for rinsing NMR tubes was stored in a pre-baked (540°C, 2h) 1000mL amber Boston round bottle with an acid cleaned (10%HCl) PTFE lined cap. The same Milli-Q water was used for rinsing all NMR tubes used in this study. After the Milli-Q water rinse, the NMR tube was then rinsed five times with ~0.5mL of sample seawater dispensed by a separate pre-baked (540°C/2h) 2mL pipet. Prior to sample loading, a new baked 2mL pipet was used to dispense one drop (~50µL) of deuterium oxide (Aldrich 99.9% D_2O, part#151882-10X0.6ML) into the NMR tube. Next 600µL of sample seawater was added to the NMR tube containing the one drop of D_2O to a height of 4.00±0.05cm from the bottom of the NMR tube. Maintaining this sample meniscus height is important for the water suppression method to work well.

Finally, the NMR tube was flame sealed, cooled and homogenized. The remaining sample was transferred to a new baked ampule and flame sealed as an archive in case a repeat measurement was needed. The NMR tube was stored at room temperature and in the dark prior to NMR analysis.

4.4 ^{1}H-NMR Analysis and Data Processing

Samples were analysed on the Bruker Avance III 600MHz NMR Spectrometer equipped with a cryoprobe at the University of Ottawa. We used a previously published water suppression pulse program for saline aqueous organic samples following the methods set forth by Lam and Simpson [32]. NMR data for each sample was acquired in overnight runs comprising n = 13,528 scans, or approximately 15 hours 15 minutes of analysis time.

The resulting NMR spectra was Fourier transformed using TopSpin v. 4.0.7. In TopSpin, all spectra were referenced to the tetramethylsilane (TMS) peak (often present at a chemical shift of 0.2 ppm) manually set to 0.0 ppm. The spectra were phased to ensure that the regions on either side of the water peak (4.8 ppm) and on either side of the TMS peak (0 ppm) were approximately flat. It is important to note that not all of the spectra required phasing, and those that did required very little phasing correction. Finally, the spectral region between the water and TMS peaks was baseline-corrected to ensure that the spectra always had an intensity greater than zero and the regions surrounding the TMS peak and the water peak were at equal intensity to each other.

Phased and baseline corrected data was imported into Microsoft Excel (v. 16.0.12527.21594) for normalization and integration. Here, the water peak was removed and replaced with a "background" baseline noise signal. The spectral data used to replace the water peak was found either above (greater than) 5.5 ppm or below (less than) 0 ppm. Since each sample spectra contained 65,536 intensity data points spanning a chemical shift range from 15 to -5 ppm,

the data density of the spectra was both reduced and "smoothed" by taking the average of every 10 data points. The TMS peak was then shifted to align with 0 ppm by finding the chemical shift value of the TMS peak and subtracting this constant value from all of the chemical shift data. Only data between 5.0 ppm and 0.2 ppm was used in our data analysis. After averaging every 10 data points and cutting out the data that did not fall in the desired chemical shift range (0.2-5.0 ppm) there were 1430 data bins. The condensed data was then normalized by dividing each intensity data bin by the sum of all the intensities in that spectrum.

4.5 Spectral Integration of Compound Classes and Statistical Analysis

Since ^{1}H-NMR is quantitative (i.e. the relative intensity of the chemical shift spectral regions indicates the number of protons of a given functional group), we can integrate spectral areas of interest in order to evaluate the percent relative abundance of unique compound classes. Examples of these compound classes included CRAM, CHO-carbohydrates, etc. The chemical shift (dH) ppm ranges of key compound classes were identified using a modified version of a table presented by Fox et al.[63] and a principal component analysis (described below). Here we provided an adapted table and figure from Fox et al.[63], 2018 (Appendix B (Table 1); Appendix A (Figure 3)) for clarity of understanding our integration and definition of key compound classes. To aid in our identification of the major compound classes present in our samples we performed a multivariate principal component analysis (PCA) in MatLab (v. 9.9.0.1538559) using the normalized data for all surface samples to identify which compound classes had the greatest impact on DOM compositional change in Baffin Bay.

We used the rule of trapezoidal elements to integrate the NMR spectral areas and for functional groups of interest. The trapezoidal elements were calculated in Microsoft Excel (v. 16.0.12527.21594) from the normalized intensity and chemical shift data using Equation 1 (T_n –

trapezoidal element of data point n, I_n – intensity of data point n, P_n – chemical shift value of data point n).

(1) $$T_n = \frac{(I_{n+1}+I_n)*(P_n-P_{n+1})}{2}$$

The sum of all of the intensities for a specific sample was then calculated (S_{tot}), as well as the sum of all of the intensities of a specific compound class for that sample (S_x). The percent relative abundance (A_x) was then calculated using Equation 2.

(2) $$A_x = \frac{S_x}{S_{tot}} * 100\%$$

Once the percent relative abundance of each compound class was calculated, this value was divided by the H/C ratio[64] ($H{:}C_x$) for that compound class to find the abundance of carbon ($C_{abundance}$). The abundance of carbon was calculated using equation 3.

(3) $$C_{abundance} = \frac{A_x}{H{:}C_x}$$

Next, the abundance of carbon atoms was multiplied by the concentration of dissolved organic carbon ([DOC]) measured from each sample to find the absolute concentration [Compound] of each compound class in all 91 samples. The absolute concentration was calculated using equation 4.

(4) $$[Compound] = C_{abundance} * [DOC]$$

This process was repeated for all compound classes identified in the samples for all 91 samples as well as for five duplicate samples which were used to ensure a low standard deviation (Appendix B (Table 2)). Duplicate samples were taken from a second ampule that was extracted from the same niskin bottle at sea. The duplicate ampule was prepared in the lab for NMR analysis

separately from the original and ran on the Bruker Avance III 600MHz NMR Spectrometer on a different day, weeks or months after the original was run. All relevant sample data can be found in Appendix B (Table 3).

Finally, Ocean Data View (v. 5.5.1) software was used to create section plots (Figure 2 and Appendix A (Figure 1)) following the two major current systems in Baffin Bay (West Greenland Current and Baffin Island Current). The colour maps depict how the concentration and abundance of DOC, CRAM, and TCHO change with depth and latitude.

5. Results and Discussion

Baffin Bay is a region of high seasonal primary productivity, with a host of diverse water masses, nutrient sources[65] and long deep water residence times (360-690 years)[66]. Thus, changes in DOM composition, and specifically the distribution and differential cycling of LDOM vs. RDOM components (e.g. TCHO and CRAM) within Baffin Bay is expected. DOC concentrations (Figure 2 A, D) ranged from 36.6 to 64.8 µmol kg^{-1} with higher average concentrations in the surface 100m (55.2 ± 5.4 µmol kg^{-1}; n = 43) vs. at depths over 200m (45.5 ± 3.5 µmol kg^{-1}; n = 36). Total ^{1}H-NMR integrations of TCHO regions and CRAM (*see methods*) indicate TCHO comprises between 7-19% and CRAM comprises between 37-59% of DOM in Baffin Bay (Figure 2). Our %TCHO estimates are on the same order of magnitude as mean %TCHO yields previously reported for North Atlantic water masses using the spectrophotometric MBTH method (3-methyl-2-benzothiazoline hydrazone; a colourimetric measurements of sugar aldehydes; 12-20%)[8,44]. From our integrations (see methods), we observe several specific DOM biomarkers are driving composition variability within Baffin Bay (e.g. CH_3-deoxysugars, protein/peptides, methanethiol,

and N-acetyl amino sugars). This work is beyond the scope of the current study and will be the subject of a follow-up paper.

These first %CRAM estimates for total seawater DOM are both surprising and unexpected. Whereas several aforementioned UF and SPE studies estimated CRAM to comprise at most 8-28% of the DOM pool[48,51], our results suggest the total CRAM pool could be up to 7 times more abundant than previously recognized. These data suggest a significant fraction of total CRAM is not retained by either UF or SPE methods. This "missing" CRAM has also eluded detection by traditional isolation and measurement (e.g. high-resolution mass spectrometry) techniques. Hertkorn et al.[51] found that 51-56% of SPE-DOM was CRAM. If we assume ~40% as a nominal SPE-DOC recovery, SPE-DOM CRAM would comprise between 20-22% of total DOC. HMW DOM CRAM was found to contain 8% of total DOC[48]. Thus, the fraction of CRAM that can be quantified using traditional isolation techniques comprises between 28-30% of the total DOC pool. Our total DOM ^{1}H-NMR results suggest CRAM comprises between 37-59% of total DOC, suggesting a significant fraction of total CRAM (up to 31%) has evaded detection by coupling these earlier techniques. This is consistent with recent work showing that SPE has a lower affinity for isolating marine humic substances due to an inherent SPE-DOM molecular composition bias[67]. If similarly high %CRAM is observed in the Atlantic and Pacific Oceans, our results would suggest CRAM is the most abundant, identifiable compound-class in the DOM pool.

Percent carbon yield estimates (e.g. %TCHO or %CRAM as % total DOC) are commonly used to evaluate broadscale changes in DOM composition, and thus to infer differential cycling of a compound-class relative to the total DOM pool[8,48,68]. Despite some station and depth specific variance, and wide overall ranges in, %TCHO and %CRAM (Appendix A (Figure 1 and 2)), we do not observe statistically significant differences between average %yields in the surface (<100m;

%TCHO = 13 ± 2% and %CRAM = 50 ± 4%; *n* = 43) and deep waters (>200m: %TCHO = 13 ± 2% and %CRAM = 51 ± 4%; *n* = 36). A one-way ANOVA test comparing surface versus deep % abundance values confirms the similarity of surface versus deep %TCHO ($F(1, 87)=0.11$, $p=0.74$) and %CRAM [$F(1, 18) =0.0004$, $p=0.98$] values. Similarly, no statistically significant average difference in surface %TCHO ($F(1, 38)=2.87$, =0.10), deep %TCHO ($F(1, 41)=0.98$, =0.33), surface %CRAM ($F(1, 38)=0.31$, $p=0.58$) or deep %CRAM ($F(1, 41)=1.93$, $p=0.17$) was observed between the two transects. Heterogeneous %TCHO distributions at depth are difficult to explain, but could represent fresh carbohydrate contributions via hydrolysis of sinking particles, resuspension from local shelf/slope regions or recent water mass ventilation. We note that similarly homogeneous %TCHO depth profiles were observed in North Atlantic waters above 40°N (%TCHO = 18-25%) together with relatively uniform DOC concentrations[44]. Our %TCHO estimates do not distinguish between poly- and monosaccharides, and thus do not preclude the microbial transformation of surface polysaccharides to more degraded monosaccharides at depth as has been previously inferred from %TCHO profiles[8]. The lack of clear broad scale trends in %TCHO and %CRAM suggests these proxies for DOM compositional change are perhaps less useful than compound-class carbon concentrations (e.g. μmol kg^{-1}) for understanding DOM biogeochemical cycling in this region.

In stark contrast to our %yield data, average CRAM and TCHO concentrations clearly decrease with depth from 26.9 μmol C kg^{-1} and 6.9 μmol C kg^{-1} above 200m (*n* = 43) respectively, to 23.9 μmol C kg^{-1} and 5.9 μmol C kg^{-1} below 200m (*n* = 43) (Figure 2). A one-way ANOVA test comparing the surface versus deep concentration values confirms the statistically significant difference of the surface versus deep TCHO ($F(1, 78)=25.29$, $p=0.0000028$) and CRAM ($F(1, 78)=36.46$, $p=0.000000041$). Carbohydrate concentrations estimated via our ^{1}H-NMR approach

are similar to those measured directly using the MBTH method[8,44], suggesting the technique closely approximates seawater TCHO concentrations. On average, we observe a 1.0 µmol C kg^{-1} loss of TCHO with depth. This is consistent with previous work in the North Atlantic showing a 3.0 µmol C kg^{-1} loss of TCHO with depth using the MBTH method[8]. Together, this suggests a similar amount of bioavailable TCHO is rapidly cycled within the upper water column of Baffin Bay on seasonal timescales.

We observe significant decreases in CRAM (3.0 µmol C kg^{-1}) with depth in Baffin Bay – from 26.9 µmol C kg^{-1} in the upper 100m (*n*=43) to 23.9 µmol C kg^{-1} below 200m (*n*=36). Decreasing CRAM concentrations with depth generally opposes the current paradigm of marine CRAM cycling. CRAM is often inferred to have an invariant, biologically refractory chemical composition that accumulates and persists in the deep ocean as the result of slow heterotrophic microbial activity over timescales of ocean mixing[50,52–54]. While our results generally support the idea that a large reservoir of CRAM (average = 23.9 µmol C kg^{-1}; n = 36) persists in deep Baffin Bay, an average 3.0 µmol C kg^{-1} loss of surface CRAM suggests that a significant fraction of the total CRAM pool (~11% of total DOC) is produced in the surface ocean (either autotrophically or through microbial heterotrophic degradation of photosynthate C). This CRAM is biologically available and rapidly removed in the surface ocean. If instead, we subtract minimum CRAM concentrations below 200m from maximum CRAM concentrations in the upper 100m for each station sampled, we estimate that bioavailable CRAM removal could be much higher (e.g. between 5.0-15.1 µmol C kg^{-1}; 17-42%). Loss of bioavailable surface CRAM with depth is consistent with recent studies observing some bacterioplankton-specific oxidation of model CRAM proxy compounds (e.g. deoxycholate)[69] and also the selective microbial degradation of CRAM produced through the incubation of testosterone induced metabolites[70]. Based on our average loss of surface

CRAM (3.0 µmol C kg^{-1}) and average loss of surface TCHO (1.0 µmol C kg^{-1}) values, our observations suggest rapid removal of surface CRAM is three times higher than that for TCHO. Together with previous studies, our data suggest that CRAM cannot simply be considered RDOM, but instead that a "two-pool model" of labile vs. recalcitrant CRAM cycling exists in this region. While further work is needed to confirm the differential cycling of individual CRAM compounds, we hypothesize that CRAM contains a continuum of molecular forms with differing chemical composition (isomers or heteroatom abundance), ^{14}C-ages, molecular sizes and biological reactivities. Similarities in the distributions of total DOC to CRAM concentrations (given by the relatively consistent %CRAM we observe), suggest that a significant fraction of the DOM cycle in Baffin Bay is mediated by the production and removal of CRAM.

Using our CRAM concentrations together with recent hypsometry and volumetric estimates of Baffin Bay, we calculate the mass of CRAM and TCHO carbon stored within deep Baffin Bay. Jakobsson and co-workers estimated surface (0-200m) and deep (200-2450m) Baffin Bay to contain 6.67 x 10^4 and 3.69 x 10^5 km^3 volume, respectively[71]. Using average CRAM and TCHO concentrations of 23.9 µmol C kg^{-1} and 5.7 µmol C kg^{-1} within deep Baffin Bay we estimate that 106.2 Tg C of CRAM and 25.5 Tg C of TCHO are stored in deep Baffin Bay. Using previously established ^{14}C residence times, recalcitrant forms of deep CRAM and TCHO likely persist for centuries (360-690 years)[66]. Similarly, using average surface CRAM and TCHO concentrations (<200m) of 26.9 µmol C kg^{-1} and 6.8 µmol C kg^{-1}, we estimate 21.6 Tg C of CRAM and 5.5 Tg C of TCHO in the surface of Baffin Bay. Subtracting the difference of surface vs. deep CRAM and TCHO concentrations would thus imply the semi-annual removal of 2.4 Tg C and 0.8 Tg C of biologically labile CRAM and TCHO molecular populations.

6. Baffin Bay DOM Molecular Composition: Summary and Implications

Taken together, our total seawater ^{1}H-NMR results are consistent with recent work challenging the notion that CRAM and TCHO can be simply considered canonical proxies for RDOM and LDOM, at least for the Baffin Bay region. We report the highest total DOC CRAM concentrations to date – suggesting that, as a compound-class, CRAM (37-59%) comprises the largest identifiable fraction of the DOM pool that has evaded other methods of detection (e.g. ultrafiltration and SPE-DOM techniques). Surface to deep gradients in CRAM occur largely in concert with changes in DOC concentrations, suggesting a large portion of the marine DOM cycle is mediated through both autotrophic production and heterotrophic cycling of this population of chemically-distinct molecules. While more work is needed to constrain CRAM distributions and cycling in the global ocean, our study suggests that up to 21-43% of surface CRAM is semi-labile material, at least within high-productivity Arctic ecosystems. These observations change our current understanding of the DOM cycle in regions of rapid Arctic climate and ecosystem change. While significant loss of bioavailable CRAM and TCHO is observed, the long-term storage of recalcitrant CRAM and TCHO in deep Baffin Bay, with centennial timescales of deep water ventilation, suggest that this material is likely to have high ^{14}C-ages and become increasingly degraded through slow microbial heterotrophy before entering global deep water circulation pathways in the North Atlantic. Future temporal studies focused on assessing changes in CRAM and TCHO will be paramount in constraining Arctic DOM carbon budgets, and assessing how the Arctic DOM reservoir might respond in the face of anthropogenic climate change.

7. Chapter Acknowledgements

We gratefully acknowledge Amundsen Science staff, Chief Scientist Alexandre Forest, Anissa Merzouk and the crew of the CCGS *Amundsen* for their support and the opportunity to participate in the 2019 research expedition. We thank Sara Zeidan for aiding in sample collection. We acknowledge support from the ArcticNet Network of Centers of Excellence of Canada and Amundsen Science as a Major Scientific Initiative Centre. We thank the NMR facility staff Dr. Glenn Facey, Dr. Peter J. Pallister, and Dr. Vincent Morin for providing training for the ^{1}H-NMR instrument and their ongoing support while running samples. We thank Claire Normandeau and Qiang Shi of Dalhousie University for performing high temperature combustion DOC measurements. This work was supported by NSERC (Natural Sciences and Engineering Research Counsil of Canada) Discovery Grant (B.D.W), and the Canada Research Chairs program (B.D.W). This work is a contribution to ArcticNet.

8. Figures with Captions

Figure 1: Map of major current systems (Baffin Island Current and Western Greenland Current) and sample site locations in Baffin Bay. (A) Major current systems in Baffin Bay as well as inflow and outflow currents from the Canadian Arctic Archipelago and the Labrador Sea are shown by the white arrows. (B) Temperature data for Baffin Bay surface water is shown by the red and blue colour map; the locations of the sampling sites are represented by the labeled black diamonds and East and West transects indicated by white dashed lines. (C) Salinity data for Baffin Bay surface water is shown by the red and blue colour map. (Adapted from Zeidan et al., 2022)[66]

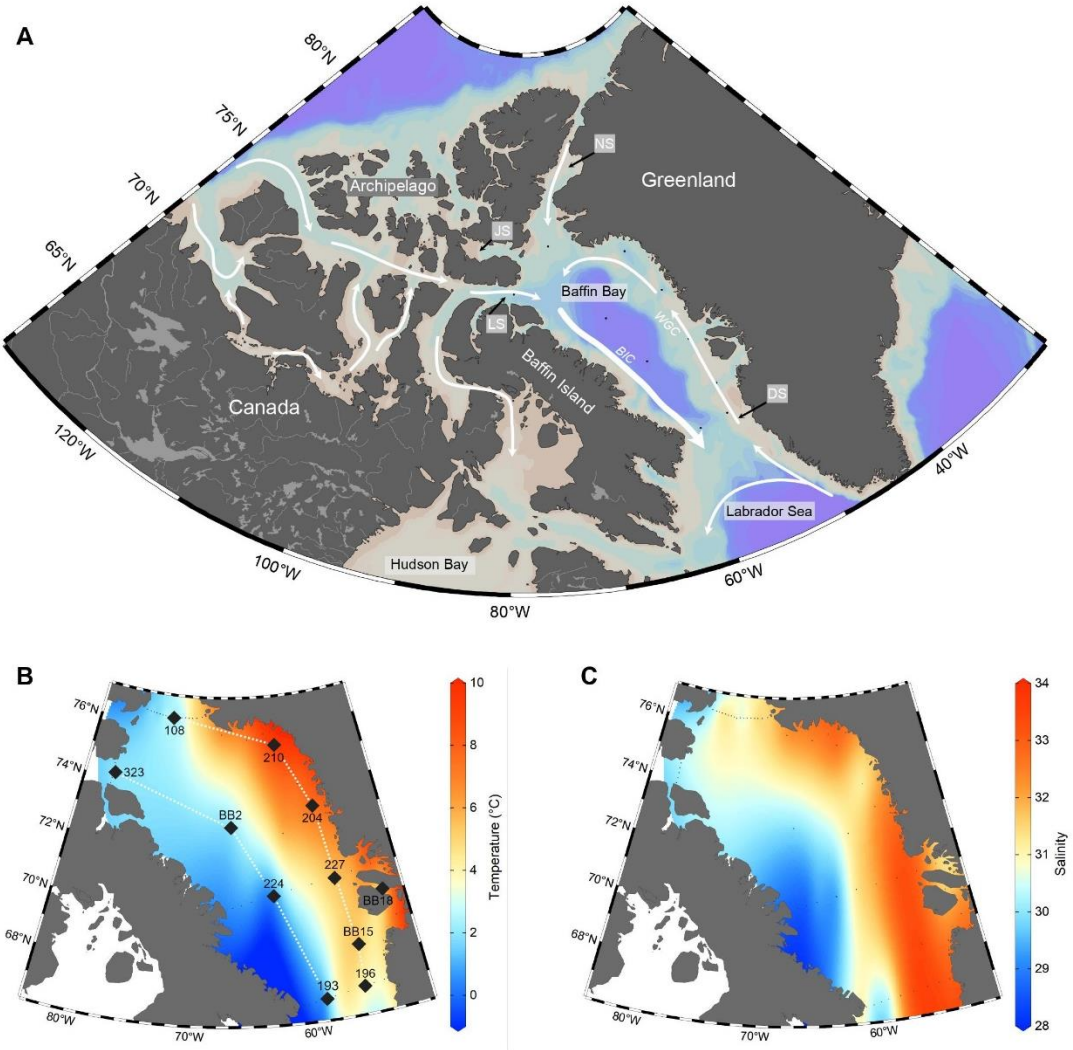
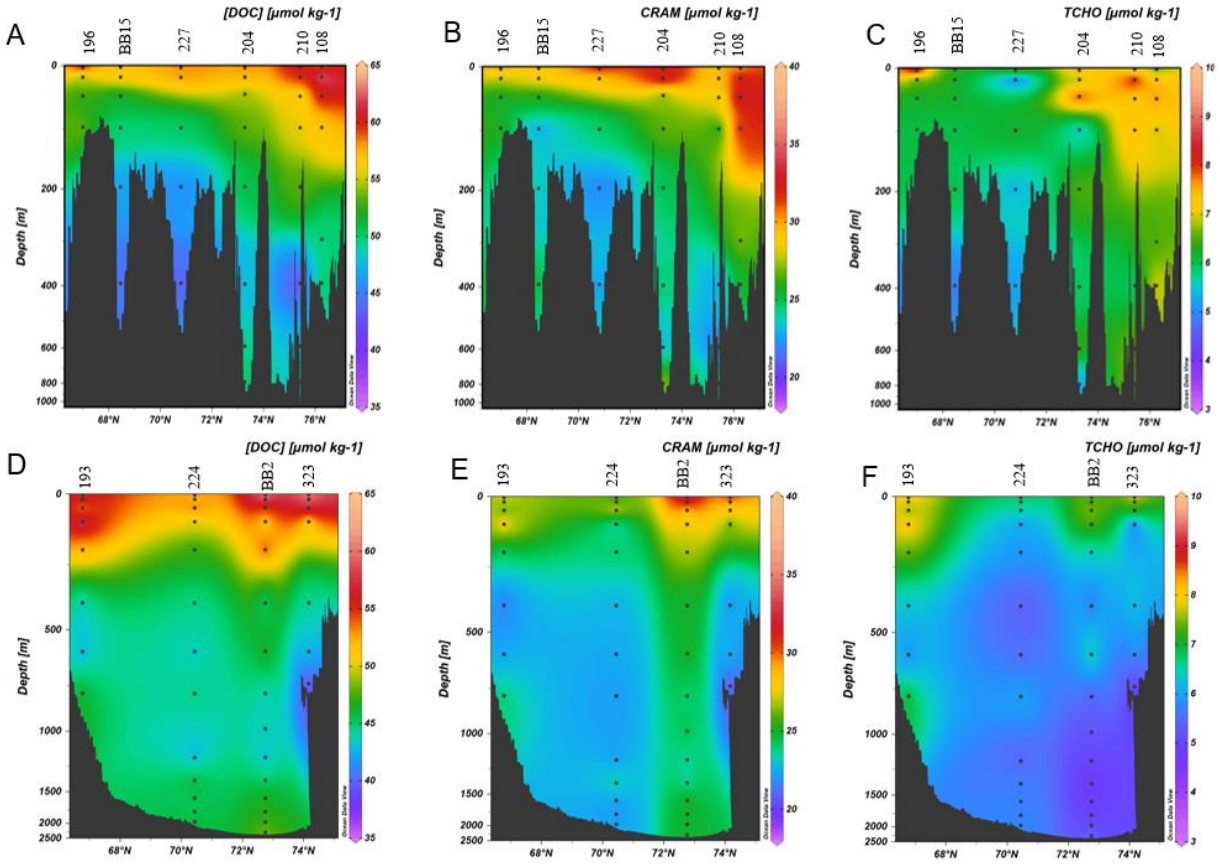


Figure 2: Section plots following the two major current systems in Baffin Bay; West Greenland Current (A, B, C) and Baffin Island Current (D, E, F). The colour map indicates the concentration of each of the compound classes; DOC (A, D), CRAM (B, E), or carbohydrates (C, F). The black circles show the location that each sample was taken. Black regions represent the bathymetry of the seafloor.



9. Conclusion

Anthropogenic climate change is impacting the carbon cycle in the Arctic Ocean; however, further measurements are required to determine the impact of global climate change on the Arctic biogeochemical cycles. Marine DOM is the largest reduced pool of actively cycling carbon in the ocean; however, despite its importance, over 80% of DOM remains uncharacterized at the molecular level[56]. Understanding the chemical structure and cycling of marine DOM is imperative to constrain the marine carbon cycle. This thesis seeks to address fundamental knowledge gaps pertaining to the chemical composition of the marine DOM pool within a rapidly changing Arctic ecosystem. Each of these aspects have their own scientific merits. First, understanding chemical changes within the total DOM reservoir without chemical- or size- bias has been the "holy grail" for marine biogeochemists for decades. The total seawater DOM composition data presented herein are the first of their kind to be reported for any ocean region. Second, given that the Arctic Ocean is warming at a rate four times faster than the global average[57], and the fact that there exists little baseline DOM composition data for the Arctic, understanding how DOM will play a role in mitigating global change is almost completely unknown. The anthropogenic acceleration of increasing temperatures, decreasing pH[72] and increasing marine productivity[73] in Baffin Bay makes it an ideal location to study the net impact of climate change on polar DOM biogeochemistry[56].

Our total seawater ^{1}H-NMR results are consistent with recent work challenging the notion that CRAM and TCHO compound classes can be simply characterized as "recalcitrant" or "labile" DOM, respectively. Using a novel integration approach and known H/C ratios for DOM functional groups, we report: i) total %DOC abundance estimates and µmol C L^{-1} concentrations of TCHO, and ii) the first CRAM µmol C L^{-1} concentrations and total %DOC abundance estimates for CRAM in seawater. TCHO comprises between 7-19% of the DOM pool however there are no clear trends between TCHO concentration and depth. These CRAM concentrations indicate that CRAM (37-59%) comprises the largest identifiable fraction of the DOM pool – a finding that was previously unknown for seawater DOM. This suggests that much CRAM in the DOM pool has evaded other isolation methods (e.g. ultrafiltration and solid phase extraction techniques). Most broadly, we find that surface to deep distributions of CRAM correlate with changes in DOC concentration, suggesting a substantial portion of the marine DOM cycle is mediated through both autotrophic

production and microbially-mediated heterotrophic degradation of this compound-class. While more work is needed to constrain CRAM distributions and cycling in the global ocean, our study suggests that up to 21-43% of surface CRAM is in fact semi-labile material, at least within this high-productivity Arctic ecosystem. We hypothesize that the rapid cycling of semi-labile CRAM and TCHO is likely occurring in other ocean regions. This hypothesis is readily testable and the magnitude of such fluxes within the DOM pool should be the focus of future global hydrographic studies.

These observations not only change our current understanding of the DOM cycle; they also add relevance to understanding mechanisms for how DOM can act as a large carbon "capacitor" for deep C storage within rapidly changing Arctic ecosystems. Using recently published hypsometry data, together with our CRAM and TCHO concentration data, we estimate that approximately 21.6 TgC of CRAM and 5.5 TgC of TCHO are present in the surface of Baffin Bay and approximately 106.2 Tg C of CRAM and 25.5 Tg C of TCHO are stored as recalcitrant DOM molecules in deep Baffin Bay. Subtracting these surface vs. deep TgC reservoirs suggests that there is a semi-annual removal of 2.4 Tg C and 0.8 Tg C of biologically labile CRAM and TCHO in Baffin Bay, respectively. The loss of surface bioavailable CRAM and TCHO is significant in shaping our understanding of marine primary production and rates of microbial respiration. For example, it places the respiration of DOM to CO_2 and/or the molecular transformation DOM from a labile to a recalcitrant chemical species (e.g. via the microbial carbon pump) into a more quantitative lens that, if observed globally, could later be compared to rates of organic matter burial, deep sea chemoautotrophy, riverine DOM fluxes, etc.

Conversely, the long-term storage of recalcitrant CRAM and TCHO in deep Baffin Bay is also of significant interest as it would be this DOM that would mitigate (store) atmospheric CO_2 on centennial timescales. Our results suggest that several orders of magnitude more CRAM and TCHO carbon is stored at depth in Baffin Bay. If truly recalcitrant material (e.g. that produced via the microbial carbon pump), we hypothesize this deep CRAM and TCHO in deep Baffin Bay to have high ^{14}C-ages and to be comprised highly degraded forms of CRAM and TCHO molecules. Using ^{14}C residence times established by Zeidan et al., recalcitrant forms of deep CRAM and TCHO likely persist for centuries (360-690 years)[66] before slowly entering global deep water circulation pathways in the North Atlantic. The Arctic Ocean carbon cycle has the potential to

influence both local and global climate systems. Future temporal studies focused on assessing changes in CRAM and TCHO will be paramount in constraining Arctic DOM carbon budgets, and assessing how the Arctic DOM reservoir might respond in the face of anthropogenic climate change.

10. References

1. Hansell, D., Carlson, C., Repeta, D. & Schlitzer, R. Dissolved Organic Matter in the Ocean: A Controversy Stimulates New Insights. *Oceanography* 22, 202–211 (2009).

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

3. Druffel, E. R. M., Griffin, S., Coppola, A. I. & Walker, B. D. Radiocarbon in dissolved organic carbon of the Atlantic Ocean. *Geophys Res Lett* 43, 5279–5286 (2016).

4. Beaupré, S. R. & Druffel, E. R. M. Constraining the propagation of bomb-radiocarbon through the dissolved organic carbon (DOC) pool in the northeast Pacific Ocean. *Deep Sea Res Part Oceanogr Res Pap* 56, 1717–1726 (2009).

5. Benner, R. & Amon, R. M. W. The Size-Reactivity Continuum of Major Bioelements in the Ocean. *Annu Rev Mar Sci* 7, 1–21 (2015).

6. Walker, B. D., Beaupré, S. R., Guilderson, T. P., McCarthy, M. D. & Druffel, E. R. M. Pacific carbon cycling constrained by organic matter size, age and composition relationships. *Nat Geosci* 9, 888–891 (2016).

7. Mitschke, N., Vemulapalli, S. P. B. & Dittmar, T. NMR spectroscopy of dissolved organic matter: a review. *Environ Chem Lett* 21, 689–723 (2023).

8. Benner, R., Pakulski, J. D., McCarthy, M., Hedges, J. I. & Hatcher, P. G. Bulk Chemical Characteristics of Dissolved Organic Matter in the Ocean. *Science* 255, 1561–1564 (1992).

9. Lam, B. & Simpson, A. J. Direct 1H NMR spectroscopy of dissolved organic matter in natural waters. *Analyst* 133, 263–269 (2007).

10. McCarthy, J. The Arctic Ocean in the 21st century with warming at a rate double the global average, the region's animal populations struggle to adapt.PDF. *The American Prospect* 19, (2008).

11. Codispoti, L. A. *et al.* Synthesis of primary production in the Arctic Ocean: III. Nitrate and phosphate based estimates of net community production. *Prog Oceanogr* 110, 126–150 (2013).

12. Hill, V. J. *et al.* Synthesis of integrated primary production in the Arctic Ocean: II. In situ and remotely sensed estimates. *Prog Oceanogr* 110, 107–125 (2013).

13. Matrai, P. A. *et al.* Synthesis of primary production in the Arctic Ocean: I. Surface waters, 1954–2007. *Prog Oceanogr* 110, 93–106 (2013).

14. Arrigo, K. R. & Dijken, G. L. van. Secular trends in Arctic Ocean net primary production. *J Geophys Res Oceans* 116, (2011).

15. Mathis, Hansell, Kadko, Bates & Cooper. Determining net dissolved organic carbon production in the hydrographically complex western Arctic Ocean. *Limnology and Oceanography* 52, 1789–1799 (2007).

16. Anderson, L. G., Jones, E. P. & Swift, J. H. Export production in the central Arctic Ocean evaluated from phosphate deficits. *J Geophys Res Oceans* 108, (2003).

17. Wheeler, P. A., Watkins, J. M. & Hansing, R. L. 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 Part Ii Top Stud Oceanogr* 44, 1571–1592 (1997).

18. Aagaard & Carmack. The Role of Sea Ice and Other Fresh Water in the Arctic Circulation. *Journal of Geophysical Research* 94, 14485–14498 (1989).

19. Amon, R. M. W. *et al.* Dissolved organic matter sources in large Arctic rivers. *Geochim Cosmochim Ac* 94, 217–237 (2012).

20. Gordeev, V. A Reassessment of the Eurasian River Input of Water, Sediment, Major Elements, and Nutrients to the Arctic Ocean. *American Journal of Science* 664–691 (1996).

21. Holmes, R. M. *et al.* Seasonal and Annual Fluxes of Nutrients and Organic Matter from Large Rivers to the Arctic Ocean and Surrounding Seas. *Estuaries Coasts* 35, 369–382 (2012).

22. Köhler, Meon, Gordeev, Spitzy & Amon. Dissolved organic matter (DOM) in the estuaries of Ob and Yenisei and the adjacent Kara Sea, Russia. *Proceedings in Marine Science* 281–309 (2003).

23. Alling, V. *et al.* Nonconservative behavior of dissolved organic carbon across the Laptev and East Siberian seas. *Global Biogeochem Cy* 24, n/a-n/a (2010).

24. Cooper, L. W. *et al.* Linkages among runoff, dissolved organic carbon, and the stable oxygen isotope composition of seawater and other water mass indicators in the Arctic Ocean. *J Geophys Res Biogeosciences* 110, n/a-n/a (2005).

25. Letscher, R. T., Hansell, D. A. & Kadko, D. Rapid removal of terrigenous dissolved organic carbon over the Eurasian shelves of the Arctic Ocean. *Mar Chem* 123, 78–87 (2011).

26. Hansell, Kadko & Bates. Degradation of terrigenous dissolved organic carbon in the western Arctic Ocean.PDF. *Science* 858–861 (2004).

27. Benner, R., Louchouarn, P. & Amon, R. M. W. Terrigenous dissolved organic matter in the Arctic Ocean and its transport to surface and deep waters of the North Atlantic. *Global Biogeochem Cy* 19, n/a-n/a (2005).

28. Amon, R. M. W. & Benner, R. Combined neutral sugars as indicators of the diagenetic state of dissolved organic matter in the Arctic Ocean. *Deep Sea Res Part Oceanogr Res Pap* 50, 151–169 (2003).

29. Bøsrheim, K. Y. & Myklestad, S. M. Dynamics of DOC in the Norwegian Sea inferred from monthly profiles collected during 3 years at 66°N, 2°E. *Deep Sea Res Part Oceanogr Res Pap* 44, 593–601 (1997).

30. Engbrodt, R. & Kattner, G. On the biogeochemistry of dissolved carbohydrates in the Greenland Sea (Arctic). *Org Geochem* 36, 937–948 (2005).

31. Fransson, A. *et al.* The importance of shelf processes for the modification of chemical constituents in the waters of the Eurasian Arctic Ocean: implication for carbon fluxes. *Cont Shelf Res* 21, 225–242 (2001).

32. Opsahl, S., Benner, R. & Amon, R. M. W. Major flux of terrigenous dissolved organic matter through the Arctic Ocean. *Limnol Oceanogr* 44, 2017–2023 (1999).

33. Amon, R. M. W. & Meon, B. 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 (2004).

34. Mathis *et al.* The Pacific Arctic Region, Ecosystem Status and Trends in a Rapidly Changing Environment. in *Biogeochemistry of the Pacific Arctic Region* 223–268 (Springer, 2014). doi:10.1007/978-94-017-8863-2.

35. Shen, Y., Benner, R., Kaiser, K., Fichot, C. G. & Whitledge, T. E. Pan-Arctic Distribution of Bioavailable Dissolved Organic Matter and Linkages With Productivity in Ocean Margins. *Geophys Res Lett* 45, 1490–1498 (2018).

36. Catalá, T. S., Shorte, S. & Dittmar, T. Marine dissolved organic matter: a vast and unexplored molecular space. *Appl Microbiol Biotechnol* 105, 7225–7239 (2021).

37. Jiao, N. Microbial production of recalcitrant dissolved organic matter: Long-term carbon storage in the Global Ocean. *Nature Reviews Microbiology* 8, 593–599 (2010).

38. Lehmann, T. E. Nuclear Magnetic Resonance Spectroscopy. *Magnetochemistry* 4, 20 (2018).

39. McMurry, J. *Organic Chemistry*. vol. xxiii (Belmont, Calif. : Brooks/Cole, 2012).

40. McCarthy, M. D., Hedges, J. I. & Benner, R. The chemical composition of dissolved organic matter in seawater. *Chem Geol* 107, 503–507 (1993).

41. Aluwihare, L. I., Repeta, D. J. & Chen, R. F. A major biopolymeric component to dissolved organic carbon in surface sea water. *Nature* 387, 166–169 (1997).

42. Aluwihare, Repeta, D., Pantoja, S. & Johnson, C. Two Chemically Distinct Pools of Organic Nitrogen Accumulate in the Ocean. *Science* 308, 1007–1010 (2005).

43. Panagiotopoulos, C., Repeta, D. J., Mathieu, L., Rontani, J.-F. & Sempéré, R. Molecular level characterization of methyl sugars in marine high molecular weight dissolved organic matter. *Mar Chem* 154, 34–45 (2013).

44. Goldberg, S. J., Carlson, C. A., Bock, B., Nelson, N. B. & Siegel, D. A. Meridional variability in dissolved organic matter stocks and diagenetic state within the euphotic and mesopelagic zone of the North Atlantic subtropical gyre. *Mar Chem* 119, 9–21 (2010).

45. Hedges, J. I., Hatcher, P. G., Ertel, J. R. & Meyers-Schulte, K. J. A comparison of dissolved humic substances from seawater with Amazon River counterparts by 13C-NMR spectrometry. *Geochim Cosmochim Ac* 56, 1753–1757 (1992).

46. Williams, P. M. & Druffel, E. R. M. Radiocarbon in dissolved organic matter in the central North Pacific Ocean. *Nature* 330, 246–248 (1987).

47. Aluwihare, L. I., Repeta, D. J. & Chen, R. F. Chemical composition and cycling of dissolved organic matter in the Mid-Atlantic Bight. *Deep Sea Res Part Ii Top Stud Oceanogr* 49, 4421–4437 (2002).

48. Hertkorn, N. *et al.* Characterization of a major refractory component of marine dissolved organic matter. *Geochim Cosmochim Ac* 70, 2990–3010 (2006).

49. Cai, R. & Jiao, N. Recalcitrant dissolved organic matter and its major production and removal processes in the ocean. *Deep Sea Res Part Oceanogr Res Pap* 191, 103922 (2023).

50. Zheng, X. *et al.* Experimental Insight into the Enigmatic Persistence of Marine Refractory Dissolved Organic Matter. *Environ Sci Technol* 56, 17420–17429 (2022).

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

52. Broek, T. A. B. *et al.* Low Molecular Weight Dissolved Organic Carbon: Aging, Compositional Changes, and Selective Utilization During Global Ocean Circulation. *Global Biogeochem Cy* 34, (2020).

53. Broek, T. A. B., Walker, B. D., Guilderson, T. P. & McCarthy, M. D. Coupled ultrafiltration and solid phase extraction approach for the targeted study of semi-labile high molecular weight and refractory low molecular weight dissolved organic matter. *Mar Chem* 194, 146–157 (2017).

54. Zigah, P. K., Minor, E. C., McNichol, A. P., Xu, L. & Werne, J. P. Constraining the sources and cycling of dissolved organic carbon in a large oligotrophic lake using radiocarbon analyses. *Geochim Cosmochim Ac* 208, 102–118 (2017).

55. Lewis, C. B., Walker, B. D. & Druffel, E. R. M. New Radiocarbon Constraints on the Global Cycling of Solid-Phase Extractable Dissolved Organic Carbon. *Geophys Res Lett* 48, (2021).

56. Benner, R. Chemical Composition and Reactivity. *Biogeochemistry of Marine Dissolved Organic Matter* 59–90 (2002) doi:10.1016/b978-012323841-2/50005-1.

57. Rantanen, M. *et al.* The Arctic has warmed nearly four times faster than the globe since 1979. *Commun Earth Environ* 3, 168 (2022).

58. Dadaglio, Dinasquet, Obernosterer & Joux, I. &. Differential responses of bacteria to diatom-derived dissolved organic matter in the Arctic Ocean. *Aquatic Microbial Ecology* 82, 59–72 (2018).

59. Ziolkowski, L. A. & Druffel, E. R. M. Aged black carbon identified in marine dissolved organic carbon. *Geophys Res Lett* 37, n/a-n/a (2010).

60. Arakawa, N. *et al.* Carotenoids are the likely precursor of a significant fraction of marine dissolved organic matter. *Sci Adv* 3, e1602976 (2017).

61. Romera-Castillo, C., Sarmento, H., Álvarez-Salgado, X. A., Gasol, J. M. & Marrasé, C. Net Production and Consumption of Fluorescent Colored Dissolved Organic Matter by Natural Bacterial Assemblages Growing on Marine Phytoplankton Exudates. *Appl Environ Microb* 77, 7490–7498 (2011).

62. Abdulla, H. A., Burdige, D. J. & Komada, T. Accumulation of deaminated peptides in anoxic sediments of Santa Barbara Basin. *Geochim Cosmochim Ac* 223, 245–258 (2018).

63. Fox, C. A., Abdulla, H. A., Burdige, D. J., Lewicki, J. P. & Komada, T. Composition of Dissolved Organic Matter in Pore Waters of Anoxic Marine Sediments Analyzed by 1H Nuclear Magnetic Resonance Spectroscopy. *Frontiers Mar Sci* 5, 172 (2018).

64. Anderson, L. A. On the hydrogen and oxygen content of marine phytoplankton. *Deep Sea Res Part Oceanogr Res Pap* 42, 1675–1680 (1995).

65. Lehmann, N. *et al.* Remote Western Arctic Nutrients Fuel Remineralization in Deep Baffin Bay. *Global Biogeochem Cy* 33, 649–667 (2019).

66. Zeidan, S. *et al.* Using Radiocarbon Measurements of Dissolved Inorganic Carbon to Determine a Revised Residence Time for Deep Baffin Bay. *Frontiers Mar Sci* 9, 845536 (2022).

67. Dulaquais, G. *et al.* Size exclusion chromatography and stable carbon isotopes reveal the limitations of solid phase extraction with PPL to capture autochthonous DOM production. *Mar Chem* 249, 104213 (2023).

68. Goldberg, S. J., Carlson, C. A., Hansell, D. A., Nelson, N. B. & Siegel, D. A. Temporal dynamics of dissolved combined neutral sugars and the quality of dissolved organic matter in the Northwestern Sargasso Sea. *Deep Sea Res Part Oceanogr Res Pap* 56, 672–685 (2009).

69. Liu, S. *et al.* Different carboxyl-rich alicyclic molecules proxy compounds select distinct bacterioplankton for oxidation of dissolved organic matter in the mesopelagic Sargasso Sea. *Limnol Oceanogr* 65, 1532–1553 (2020).

70. Liu, Z. *et al.* Direct Production of Bio-Recalcitrant Carboxyl-Rich Alicyclic Molecules Evidenced in a Bacterium-Induced Steroid Degradation Experiment. *Microbiol Spectr* e04693-22 (2023) doi:10.1128/spectrum.04693-22.

71. Jakobsson, M. Hypsometry and volume of the Arctic Ocean and its constituent seas. *Geochem Geophys Geosystems* 3, 1–18 (2002).

72. Burgers, T. M., Tremblay, J.-É., Else, B. G. T. & Papakyriakou, T. N. Estimates of net community production from multiple approaches surrounding the spring ice-edge bloom in Baffin Bay. *Elementa Sci Anthropocene* 8, (2020).

73. Lewis, K. M., Dijken, G. L. van & Arrigo, K. R. Changes in phytoplankton concentration now drive increased Arctic Ocean primary production. *Science* 369, 198–202 (2020).

74. Quan, T. M. & Repeta, D. J. Periodate oxidation of marine high molecular weight dissolved organic matter: Evidence for a major contribution from 6-deoxy- and methyl sugars. *Mar Chem* 105, 183–193 (2007).

Appendix A: Supplemental Figures with Captions

Figure 1: Section plots following the two major current systems in Baffin Bay; West Greenland Current (A, B) and Baffin Island Current (C, D). The colour map indicates the percent abundance of each of the compound classes; CRAM (A, C), or carbohydrates (B, D). The black circles show the location where each sample was taken. Black regions represent the bathymetry of the seafloor.

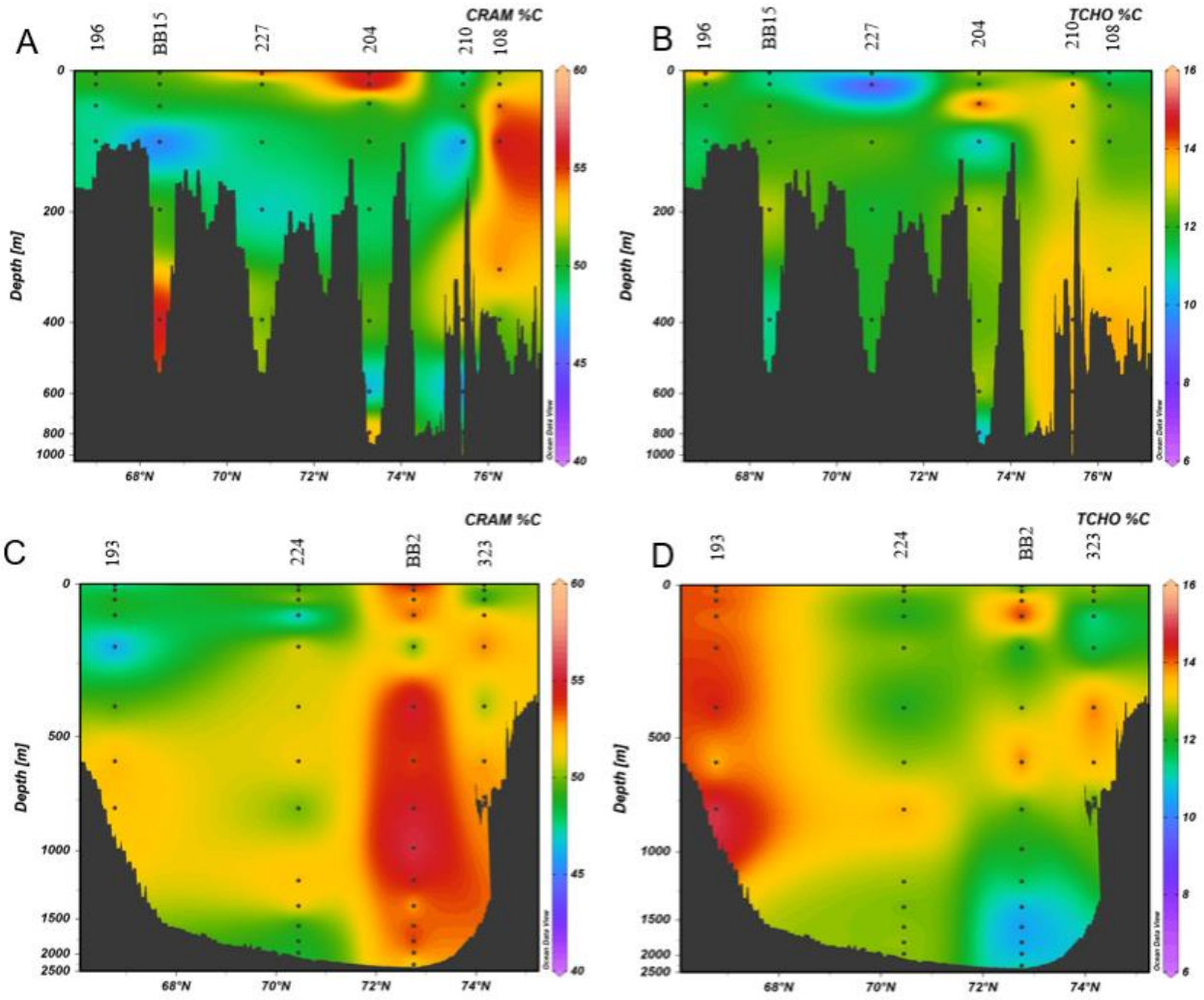


Figure 2: Depth profiles from all stations. (A, B) Percent relative abundance of CRAM and TCHO respectively. (C, D) Concentration of CRAM and TCHO respectively.

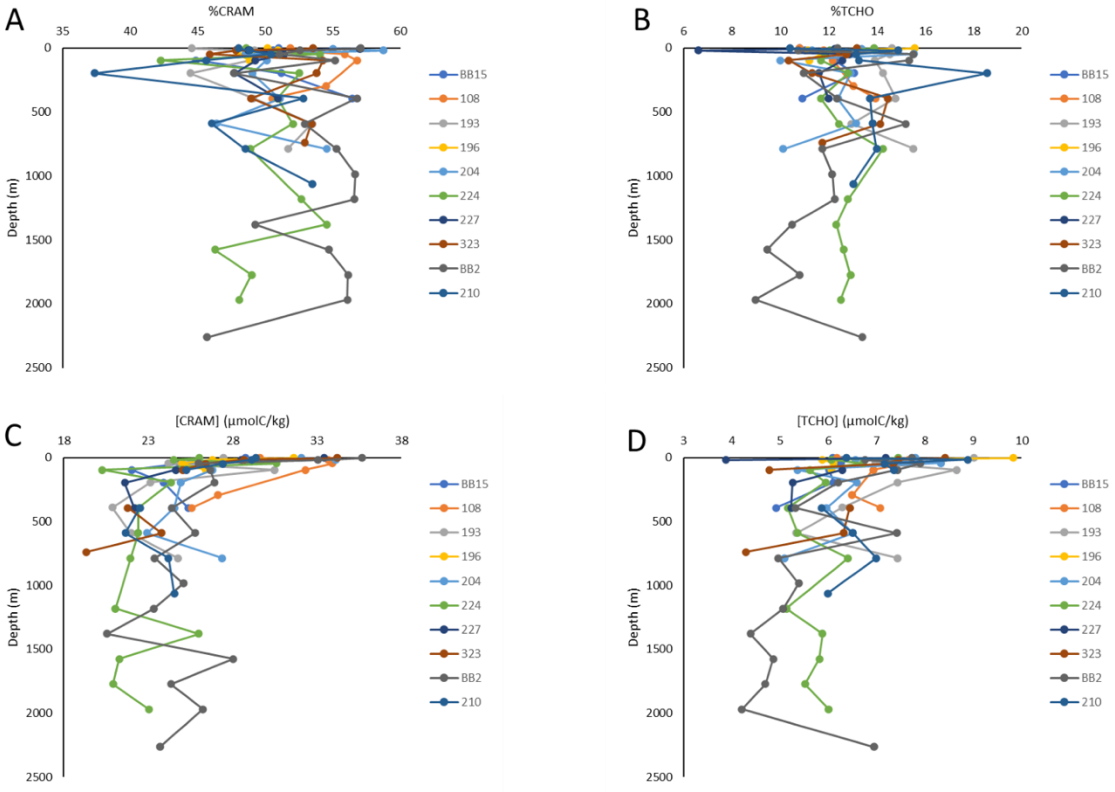
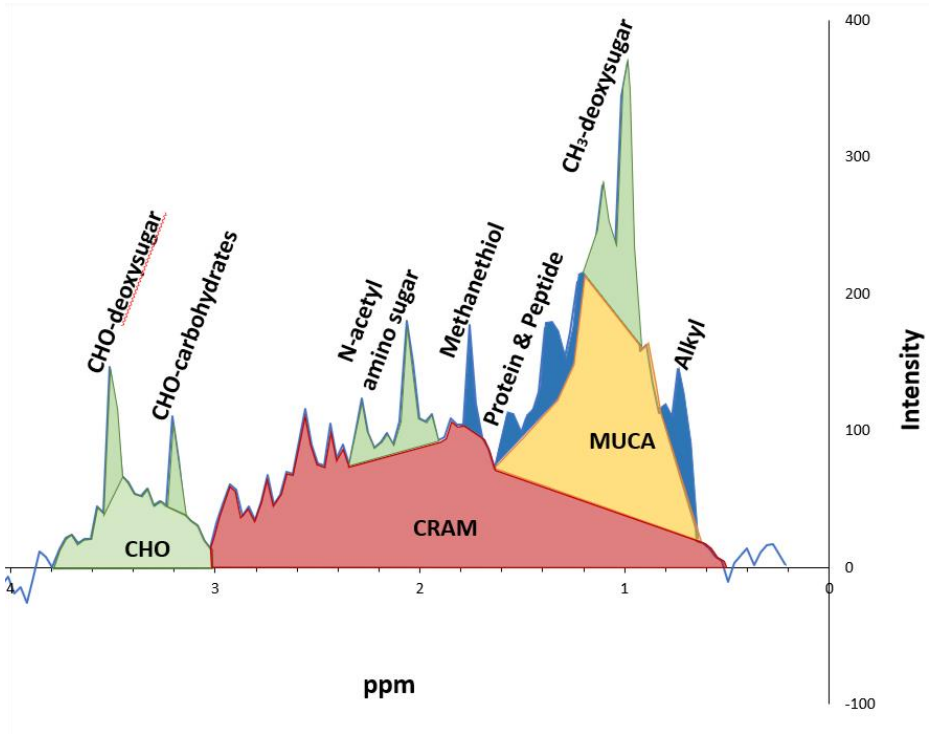


Figure 3: Example spectrum with shaded regions showing the integration areas of each of the major compound classes. Major compound classes identified via principal component analysis of all 91 samples collected.



Appendix B: Supplemental and Extended Data Tables with Captions

Table 1: List of major ^{1}H NMR compound classes and corresponding chemical shift (d_H) ppm values for seawater DOM. The ppm range corresponds to the chemical shift of the measured proton. The assignment indicates the position of the measured proton within the compounds functional group. The compound class is a broad classification that encompasses many types of similar molecules. This table was adapted from sediment porewater DOM values following Fox et al., 2018.

d_H (ppm) [c]	H:C	**H**-Assignment	Compound Class [c]
0.68-0.84	3.00	C**H**$_3$X	Alkyl (Methyl)
0.84-1.24	1.67[a]	C**H**$_3$ in 6-deoxysugars[d]	Carbohydrates
1.24-1.51	1.58 [a]	C**H**$_2$-C-CO(NHR) [c]	Protein and peptides
1.70-1.84	2.00 [a]	C**H**$_3$-C-SH [c]	Methanethiol
1.95-2.22	1.67 [a]	C**H**$_3$-C=O-NH[e]	N-acetyl amino sugars; acetate derivatives
2.72-2.80	1.67 [a]	C**H**$_3$-CH$_2$-C=O-NH [c]	Amino sugars
2.94-3.15	1.58 [a]	C**H**$_2$-NH$_2$ [c]	Protein and peptides
1.51-1.64	1.00[b]	**H**C(C)-COX [b]	Subsection of CRAM
2.22-2.72	1.00 [b]	**H**C(C)-COX [b]	Subsection of CRAM
0.91-3.15	1.00 [b]	**H**C(C)-COX [b]	CRAM
3.15-4.14	1.67 [a]	**H**C-O [c]	Carbohydrates

[a]H/C ratios of major biochemical classes of Anderson (1995) [64] were assigned

[b]H/C ratio and H-assignment data adapted from Hertkorn (2013)[51]

[c]Fox (2018)[63]

[d]Quan and Repeta (2007)[74]

[e]Aluwihare et al. (2005)[42]; Hertkorn et al. (2013)[51]

Table 2: Abundance (%) and Concentration (µmol C kg^{-1}) of major compound classes as well as standard deviation (SD) in duplicate (DUP) samples. Duplicates were extracted for NMR analysis from separate ampules. Dates of NMR preparation and analysis are listed below each duplicate number as yyyy/mm/dd.

			Alkyl (methyl) ($R-CH_3$)	**CH_3-deoxysugar**	**CH_3-deoxysugar**	**Protein & peptide**	**Amine / peptide**	**Protein & peptide**	**Methanethiol**	**MUCA**	**CHO-deoxysugar (O-methylrhamnose / O-methylfucose)**	**CRAM**	**CHO-carbohydrates**	**TCHO**
		ppm	0.68-0.84	0.84-1.05	1.05-1.20	1.20-1.31	1.31-1.45	1.45-1.65	1.71-1.79	0.68-1.65	3.44-3.56	0.65-3.03	3.03-3.81	0.93-1.20, 1.86-2.34, 3.03-3.81
stn193 100m	Abundance (%)	DUP1 2020/06/09	0.35	1.21	1.25	0.58	0.69	0.47	0.39	26.37	1.46	49.26	8.67	13.20
		DUP2 2020/10/14	0.36	0.60	1.18	0.22	0.66	0.27	0.37	23.47	1.36	48.98	10.12	13.90
		SD	**0.01**	**0.43**	**0.05**	**0.26**	**0.02**	**0.14**	**0.01**	**2.05**	**0.07**	**0.20**	**1.02**	**0.49**
	Concentration (µmol kg^{-1})	DUP1 2020/06/09	0.22	0.75	0.78	0.36	0.43	0.29	0.24	16.41	0.91	30.65	5.40	8.22
		DUP2 2020/10/14	0.23	0.37	0.74	0.13	0.41	0.17	0.23	14.60	0.84	30.48	6.30	8.65
		SD	**0.01**	**0.27**	**0.03**	**0.16**	**0.01**	**0.09**	**0.01**	**1.28**	**0.20**	**0.12**	**0.09**	**0.30**
stnBB2 sfc	Abundance (%)	DUP1 2020/07/28	0.15	1.17	0.90	0.29	3.15	0.38	0.22	15.22	1.20	57.04	8.88	12.36
		DUP2 2020/03/13	0.25	0.44	1.21	0.27	0.57	0.37	0.30	25.31	0.77	52.20	10.90	13.42
		DUP3 2021/02/06	0.12	0.59	1.13	0.12	0.79	0.15	0.47	19.59	1.23	52.89	9.01	12.30
		SD	**0.07**	**0.39**	**0.16**	**0.09**	**1.43**	**0.13**	**0.13**	**5.06**	**0.26**	**2.62**	**1.13**	**0.63**
	Concentration (µmol kg^{-1})	DUP1 2020/07/28	0.09	0.73	0.56	0.18	1.97	0.23	0.14	9.51	0.75	35.66	5.55	7.73
		DUP2 2020/03/13	0.15	0.28	0.76	0.17	0.36	0.23	0.19	15.82	0.48	32.63	6.81	8.39
		DUP3 2021/02/06	0.08	0.37	0.71	0.08	0.50	0.09	0.30	12.25	0.77	33.06	5.63	7.69
		SD	**0.04**	**0.24**	**0.10**	**0.06**	**0.89**	**0.08**	**0.08**	**3.16**	**0.16**	**1.64**	**0.71**	**0.39**

Table 3: Summary of all relevant sample information measured/calculated.

Station	Latitude °N	Longitude °W	Depth (m)	Temp °C	Salinity pss-78	[DOC] [µmol kg-1]	%CRAM %C	%TCHO %C	[CRAM] [µmol kg-1]	[TCHO] [µmol kg-1]
BB15	68.4513	-55.8998	2.5	5.8	33.5	56.4	50.99	10.8	28.76	6.1
			18.0	5.1	33.5	57.1	50.55	11.0	28.86	6.3
			47.6	-0.2	33.6	49.7	50.08	12.6	24.87	6.2
			97.5	0.9	33.9	49.1	44.87	12.2	22.02	6.0
			196.0	3	34.4	46.7	51.16	13.1	23.90	6.1
			393.4	3.7	34.6	45	56.43	10.9	25.37	4.9
108	76.2586	-74.6015	2.0	3.1	31.1	57.1	51.86	10.8	29.61	6.2
			18.0	1.8	31.6	64.8	51.95	11.8	33.67	7.6
			47.6	-1.5	32.4	60.7	55.87	12.5	33.93	7.6
			97.1	-1.3	32.9	56.9	56.77	12.2	32.30	6.9
			294.6	-0.1	34.2	49.8	54.51	13.0	27.15	6.5
			393.8	-0.1	34.3	50.7	50.50	14.0	25.60	7.1
193	66.7697	-59.3402	2.2	-0.1	30.4	61.6	44.57	14.6	27.46	9.0
			18.0	2.8	32.8	55.7	49.20	13.4	27.43	7.5
			47.8	-1.3	33.2	47.9	50.51	14.5	24.21	7.0
			97.1	-0.9	33.6	62.2	48.98	13.9	30.48	8.6
			195.8	4.4	34.6	52.1	44.44	14.3	23.14	7.4
			392.6	3	34.5	42.5	49.10	14.8	20.88	6.3
			591.9	1.5	34.5	41.1	53.46	13.0	21.98	5.3
			786.8	1	34.5	47.9	51.67	15.5	24.75	7.4
196	66.9868	-56.0595	3.3	4	33.5	63.1	50.14	15.6	31.63	9.8
			18.0	2.6	33.6	52.7	50.90	11.1	26.82	5.9
			47.4	1.7	33.7	52.1	47.98	11.7	25.00	6.1
			96.3	1.4	33.7	54	48.83	11.2	26.39	6.0
204	73.2685	-57.9995	3.4	7.3	32.2	58.3	54.99	13.4	32.08	7.8
			17.7	2.8	32.8	58	58.75	11.3	34.10	6.6
			44.6	-0.8	33.4	53.6	48.54	15.5	26.00	8.3
			97.1	-0.5	33.7	53.6	50.10	10.0	26.83	5.4
			195.8	0.7	34.0	50.9	49.01	12.9	24.95	6.6
			395.6	1.5	34.4	48.1	51.06	12.4	24.55	6.0
			591.0	1.5	34.5	49.4	46.40	13.2	22.94	6.5
			788.7	1.3	34.5	50.2	54.55	10.1	27.39	5.1
224	70.4407	-62.9815	2.3	0	29.7	53.6	48.60	13.9	26.04	7.4
			18.1	-1.4	32.4	50.2	48.84	12.0	24.54	6.1
			47.6	-1.7	32.9	56.7	54.04	11.8	30.61	6.7
			97.2	-1.7	33.1	48	42.25	11.7	20.29	5.6
			196.1	-0.6	33.8	46.4	52.51	12.8	24.38	5.9

			394.0	1.9	34.4	44.1	50.88	11.7	22.43	5.1
			591.3	1.7	34.5	43.1	52.02	12.4	22.42	5.4
			788.6	1.3	34.5	45	48.88	14.2	21.97	6.4
			1182.3	0.5	34.5	40	52.67	12.8	21.06	5.1
			1379.2	0.1	34.5	47.7	54.54	12.3	26.01	5.9
			1575.5	-0.1	34.5	46	46.28	12.6	21.30	5.8
			1772.1	-0.3	34.5	42.7	48.97	12.9	20.92	5.5
			1968.6	-0.3	34.5	48	48.06	12.5	23.05	6.0
227	70.7968	-56.9878	3.4	5.4	33.3	58.7	56.97	12.2	33.42	7.2
			18.3	5.4	33.3	58.6	48.74	6.6	28.54	3.9
			47.8	-0.6	33.6		51.39	12.9	0.00	
			97.7	0.2	33.8	50	49.25	12.6	24.64	6.3
			196.4	1.8	34.2	45.3	47.73	11.6	21.61	5.3
			393.9	2.2	34.5	43.5	50.92	12.0	22.18	5.2
323	74.1595	-80.4705	2.1	2	30.2	63.9	53.53	13.2	34.18	8.4
			18.1	-1	32.2	60	47.79	12.4	28.67	7.4
			47.9	-1.7	32.6	57.6	45.87	12.7	26.42	7.3
			97.1	-1.3	32.9	46.2	54.24	10.3	25.05	4.8
			196.0	-1.1	33.8		53.78	11.3	0.00	
			393.4	0.9	34.3	44.6	48.94	14.5	21.80	6.4
			591.5	0.9	34.5	44.6	53.40	14.2	23.80	6.3
			739.2	0.6	34.5	36.6	52.92	11.7	19.36	4.3
BB2	72.7505	-66.9965	2.1	2.2	29.8	62.5	57.05	12.4	35.66	7.7
			18.4	-0.9	31.9	62.9	52.55	10.7	33.04	6.8
			47.6	-1.7	33.0	50.9	51.02	15.5	25.99	7.9
			97.0	-1.4	33.4	48.3	55.17	15.3	26.67	7.4
			196.0	0.2	34.0	56.5	47.68	11.0	26.94	6.2
			393.8	1.4	34.4	43	56.78	12.3	24.42	5.3
			591.2	1.4	34.4	48.8	52.90	15.2	25.79	7.4
			788.3	1.7	34.5	42.3	55.27	11.7	23.39	5.0
			985.3	1.2	34.5	44.3	56.66	12.1	25.09	5.4
			1182.0	0.7	34.5	41.3	56.58	12.3	23.35	5.1
			1379.3	0.4	34.5	41.8	49.23	10.5	20.57	4.4
			1575.3	0	34.5	51.3	54.68	9.5	28.04	4.8
			1772.0	-0.2	34.5	43.4	56.11	10.8	24.35	4.7
			1967.5	-0.3	34.5	46.8	56.07	9.0	26.24	4.2
			2261.9	-0.3	34.5	51.9	45.67	13.4	23.69	6.9
BB18	70.09	-52.74	2.1	4.4	31.1	49.9	41.97	15.1	20.95	7.6
			18.2	1.2	32.7	47.7	55.26	10.6	26.39	5.0
			47.6	0.9	33.3	45.3	46.32	15.1	20.98	6.9
			97.1	0.7	33.7	47.4	49.35	13.8	23.40	6.6

			195.7	1.3	34.0	50	53.33	10.7	26.68	5.4
			395.0	1.9	34.3	46.4	58.30	12.8	27.07	6.0
			2.2	9.5	32.9	61.2	48.00	10.4	29.40	6.4
			17.8	0	33.1	59.7	48.83	14.9	29.15	8.9
			47.8	-0.9	33.5	54.4	50.40	13.1	27.44	7.1
			97.3	-0.5	33.7	55.4	45.60	13.3	25.26	7.3
210	75.4193	-61.56	196.4	0.4	34.0	53.2	37.33		0.00	
			393.9	1.9	34.4	42.6	52.82	13.7	22.51	5.8
			590.8	1.7	34.5	47.1	46.01	13.8	21.65	6.5
			788.3	1.5	34.5	49.9	48.50	14.0	24.21	7.0
			1062.2	1.3	34.5	45.9	53.47	13.0	24.56	6.0

Appendix C: Supplemental Discussion

Additional Discussion regarding trends in % abundance and concentrations in Baffin Bay

Appendix A figure 1 shows section plots following the two major current systems in Baffin Bay; West Greenland Current (A, B) and Baffin Island Current (C, D). The colour map indicates the percent abundance of each of the compound classes; CRAM (A, C) or carbohydrates (B, D). Appendix A (figure 2) depicts depth profiles of abundance and concentration of both CRAM and TCHO. These figures show no obvious spatial trend in the abundance of the two compound classes. Given the complexity of water masses and DOM sources within Baffin Bay [33] such consistency in average %CRAM and %TCHO throughout Baffin Bay is somewhat surprising. Lower %TCHO yields could be the product of known ^{1}H-NMR water suppression method attenuation of the CHO region (3.03-3.81 ppm), but also the region in which samples were taken. A few station specific changes in %TCHO are observed that can be attributed to water mass features. For example, high concentrations of [DOC] observed within the Baffin Island Current (BIC; Stations 193, 224; 0-250m. Figure 2F and Appendix A (figure 2C)) co-occur with high %TCHO (Appendix A figure 1D), perhaps indicating export of DOM with unique chemical composition, largely modified by Pacific and Canadian Arctic Archipelago source contributions, to the Labrador Sea. High %TCHO within the nephloid layer of Station BB2 (13.4%) relative to overlying deep water (average = 10% TCHO from 1400-2000m) is a strong indication of new labile DOM being added to Baffin Bay deep water. Furthermore, Appendix A (figure 1C) shows high %CRAM at station BB2 which corresponds to the high CRAM concentrations at the same station (figure 2). Station BB2 is located in the center of the Baffin Bay gyre and is farther from the Baffin Island coast compared to the other stations in this transect. It is possible that station BB2 is located within different water masses, with their own unique local characteristics, which result in higher %CRAM and CRAM

concentrations compared the other stations in this transect. This hypothesis requires further sampling within Baffin Bay to constrain the water mass boundaries more precisely.

Appendix A (figure 3) shows an example spectrum with shaded regions depicting the integration areas of each of the major compound classes. The major compound classes were identified via a principal component analysis of all 91 samples collected. The green areas, which include CHO, CHO-deoxysugar, CHO-carbohydrates, N-acetyl amino sugar, and CH_3-deoxysugar, were amalgamated to form the area of TCHO. The area used to calculate CRAM is shaded in red. The compound classes that were not discussed in this study include molecularly uncharacterized aliphatic (MUCA) shaded in orange, as well as methanethiol, proteins & peptides, and alkyls shaded in blue.

Appendix D: MatLab Principal Component Analysis Code from the Statistics and Machine Learning Toolbox

```
>> load('Intensity.csv');
>> load('ppm.csv');
>> sfc_inv=Intensity';
>> [CF,SC,LA]=pca(sfc_inv);
>> LA_percent=100*LA/sum(LA);
>> pareto(LA_percent)
>> xlabel('Principal Component')
>> ylabel('eigen values (%)')
>> names=['BB2A'; '108 '; '323 '; '204 '; '224 '; '227 '; 'BB15'; '196 '; '193 '; 'BB18'; '210 '];
>> X1=LA_percent(1);
>> X2=LA_percent(2);
>> scatter(SC(:,1),SC(:,2),'+')
>> xlabel('PC1')
>> ylabel('PC2')
>> gname(names)
>> pca_1=SC(:,1)';
>> pca_2=SC(:,2)';
>> Lpca_1=pca_1*sfc_inv;
>> Lpca_2=pca_2*sfc_inv;
>> plot(ppm, Lpca_1)
>> xlabel('ppm')
>> ylabel('Intensity')
>> plot(ppm, Lpca_2)
>> xlabel('ppm')
>> ylabel('Intensity')
```