The chemical properties of dissolved organic matter as a function of seasonal and microbiological factors

By

Antoine Vallette Viallard

A thesis submitted to the

School of Graduate Studies

in partial fulfillment of the

requirements for the degree of

Master of Science

Environmental Science / Faculty of Science

Memorial University of Newfoundland

September 2022

St. John's Newfoundland

Abstract

Dissolved organic matter (DOM) is the pool of molecules predominantly produced from cellular growth in both terrestrial and aquatic systems and forms a reservoir of 662 Pg of carbon in the ocean. With as many as 10^5 to 10^7 different chemicals held in a single sample, the chemical diversity typically outstrips the capability of analytical techniques and the human capacity to effectively monitor the effects of environmental factors on their individual abundance. To address this issue, we adopted a "fingerprinting" approach and performed two sets of experiments to monitor the behavior of co-clustered compounds.

In the first experiment, the change of DOM under seasonal, spatial, and reactivity variables was delineated using a size-exclusion chromatography approach applying multiple detectors and a computing technique called PARAFAC. The model showed how the molar mass and fluorescent properties of DOM change with the impact of biological activity and photodegradation in terrestrial aquatic systems, as well as leaching from different soil and sediment profiles. The second experiment analyzed the production of DOM moieties during the growth of several mixed diatom assemblages. Various patterns in fluorescent molecules and NMR bands were observed characterizing better the deep biological imprint of primary producers on DOM in estuaries.

These two experiments, both performed in boreal regions, were complementary both in processes (i.e., production vs degradation) and in techniques (e.g., mass spectrometry vs NMR). It also enabled us to map those effects across the aquatic gradient (i.e., rivers and coasts). Interesting findings were yielded, and each time a limited number of factors were able to explain most of the data variance which allowed me to situate and discuss my results within the context of various DOM studies.

Acknowledgments

This work has been an extraordinary marathon that benefited daily from the intellectual input of my supervisor Dr. Heather E. Reader, Research chair in Chemistry of the Ocean and Atmosphere, and of my friend and lab mate Kavi Heerah, a Ph.D. student in Environmental Science. It could not have been possible without the financial support from NSERC (CGSM-2021 and the discovery grant, respectively attributed to me and Heather), and of the Memorial University of Newfoundland. The latter proved to hold passionate professors (Dr. Uta Passow and Dr. Rachel Sipler to name only a few) as well as dedicated associations such as GSU and TAUMUN.

My work is in part the continuation of studies by Dr. Reader, Dr. Nikoline Nielsen, Dr. Colin Stedmon, and Dr. Emma Kriztberg done during Heather's postdoctoral research in Sweden (2012-2015). Hence, additional credit should be given to them as well as to their respective institutes – the University of Copenhagen; Technical University of Denmark; and Lund University – and to their funding agencies: the Swedish Research Council, the MMSB project, and the Danish Research Council for Independent Research. Their collaboration prior to but also during my work was truly appreciated. Of course, the shared material of this thesis is published upon agreement with intellectual property owners.

In bulk, I would also like to thank Thierry, Sandrine, Thibault, Camille, Jean, Marie-Joe, Daniel, Jomana, Becky, Shefali, Patrick, Nietzsche, Stephen West, Thomas Fersen, Tania, Larissa, Sarah, Diego, and Mary for reasons they will recognize. The meaningful relations we developed were extremely supportive across those 30 months. A special thought to the 229ème Notre-dames-des-Neiges scout group which opened me to a passionate observation of my environment. For that matter, I would like to respectfully acknowledge the beautiful ecosystem and territory in which I lived during this degree, the ancestral homelands of the Beothuk.

We now share this place with what is as well the homelands of the Mi'kmaq, the Inuit of Nunatsiavut and NunatuKavut, and the Innu of Nitassinan, and I believe that a respectful relationship with all the peoples of this province is feasible, so that we can strive for a collective healing and a true reconciliation for honoring this beautiful land together.

Merci,

General summary

The processes at the center of an ecosystem like biological activity and material transport, possess a molecular imprint. More precisely, a signature is held in the composition of the naturally dissolved compounds of river water, called dissolved organic matter, or DOM. Although generations of studies learned how to interpret the optical markers of this signature, our understanding of the individual molecular characteristics remains imprecise. Therefore, we designed two experiments assessing the impact of seasonal and microbiological factors on the DOM compounds' mass distribution, reactivity to reprocessing, biochemical composition, as well as their optical properties.

While the two experiments unfold in quite different contexts with 1) Sweden freshwater samples, and 2) coastal Newfoundland planktonic blooms, we observed common behaviors indicating that those thousands of chemicals have interconnected abundances. More precisely, instead of independent fates, the proportions of several moieties covaried under the effect of a limited number of factors. Yet, the intensity of those factors is modulated by climate change and will therefore affect DOM chemical properties. Considering that the quality of organic carbon forms affects its rates of transformation, my results provide insights into the effects of climate change in the ecosystem we inhabit.

Table of content

List of Tables

List of Figures

List of Appendices

List of abbreviations and symbols

BCP – MCP	Biological carbon pump – Microbial carbon pump
BGE	Biological growth efficiency (produced biomass/consumed biomass)
BOD	Biological oxygen demand (in μmol O_2 / L)
CDOM – FDOM	Chromophoric DOM – Fluorescent DOM
CRAM	Carboxyl-rich alicyclic molecules
DFAA	Dissolved-free amino acids
DFNS	Dissolved free neutral sugars
DIN – DON	Dissolved inorganic nitrogen – Dissolved organic nitrogen
DIC – DOC	Dissolved inorganic carbon – Dissolved organic carbon
DOM – POM	Dissolved organic matter – Particulate organic matter
ED	The energy density (in kJ / g of organic matter)
ESI – MS	Electrospray Ionization – Mass Spectrometry
FTIR	Fourier transform infrared spectroscopy
LDOM – RDOM	Labile DOM – Refractory DOM
MANOVA	Multivariate analysis of variance
NPOC	Non-purgeable organic carbon
NWA	North-West Atlantic
PARAFAC	Parallel factor analysis
PPL	Bond Elut PPL (Priority PolLutant) is a styrene-divinylbenzene resin
QA/QC	Quality assurance / Quality control
SEC – HPLC	Size exclusion chromatography – High-performance liquid chromatography
SPE	Solid Phase extraction
Sr – $SUVA_{254}$	Spectral slope ratio – Specific UV absorption at 254 nm
TOC – TN	Total organic carbon (in μM) – Total nitrogen (in μM)

Chapter 1: Introduction

1.1 DOM: a complex chemical pool, dynamic across space and seasons

Dissolved organic matter (DOM) is a mixture of organic chemicals present in natural water. Biochemicals, originating mostly from the activity of saprotrophs and microbes, are released from the soil and the benthic boundary layer as well as directly produced within the water column [*Hansell and Carlson*, 2014]. This huge degree of molecular diversity is transported across the riverine ecosystems down to estuaries and the worldwide ocean. "How do environmental factors integrate to shape the chemical signature of the DOM?" is in short, the core question of this thesis. The present section provides an overview of the complexity of DOM composition as well as its variation across space and seasons.

Operationally, DOM is defined as the compounds remaining in solution after filtration at 0.2, 0.45, or 0.7 μm. This distinguishes it from the particulate fraction of a water sample (particulate organic matter, POM), which represents the non-living organic material over these cut-offs. Yet, such a broad definition of DOM provides little insight into the exact chemical nature of those thousands of DOM chemicals found in a single sample. Importantly, 0.2 μm is still far larger than individual molecular size, and this suggests that aquatic OM is more accurately pictured as hydrated organic clusters along a size continuum of DOM to POM. Along that gradient, colloids and marine gels form substructures of complex shapes, with variable flexibility, viscosity, and physicochemical properties as "dissolved" organic compounds remain partially associated in aqueous solutions [*Verdugo et al.*, 2004]. The DOM pool is therefore in constant physical equilibrium with a portion of the POM, where moieties dynamically go in and out of the pool.

As it is dissolved in natural waters, one expects DOM functional groups to be hydrophilic and even ionic to tolerate saltwater conditions [*Estrada et al.*, 2007; *Huuskonen*, 2000]. Yet, evidence shows that double bonds and conjugated systems are also common functionalities of

DOM, by cooperatively solubilizing within macromolecular frameworks [*Burkhard*, 2000]. Certain chemical groups are also affected by pH and redox conditions, which modify their charge and oxidation state. Recent DOM isolation techniques, such as solid-phase extraction (SPE), utilize sorption properties [*Chen et al.*, 2016]. For example, a commonly used sorbent called PPL requires that the pH of the sample be lowered to 2 to induce carboxylic acid protonation and adsorb DOM.

Extraction of DOM by SPE has been widely used to increase analyte concentrations while simultaneously washing out salts that interfere with measurements. Such treatments on bulk DOM facilitate precise compound identification in DOM with analytical techniques. Yet, a complete characterization of molecules in the DOM pool has been an ongoing challenge for decades with 80% to 97% of DOM remaining to be identified [*Catalá et al.*, 2021].

Our understanding is therefore quite limited, yet since most DOM is derived from biomass, studies have started by monitoring the abundance and the source of certain biochemical families [*Cooper et al.*, 2020; *Hansell and Carlson*, 2014]. DOM molecular classes are diverse, including, non-exhaustively, saccharides (e.g., glucose, cellulose, hemicellulose, peptidoglycan, etc.), amino acids (both standard and not, freely dissolved or assembled in peptides), and lipids (e.g., phospholipids, glycolipids, glycerol residues, polyketides, etc.). More complex structures range from water-soluble vitamins (i.e., vit. C, B_1 to B_{12}), polyaromatics (e.g., lignin, tannins, etc.), and nitrogenous bases (e.g., extracellular DNA, adenine cofactors, etc.). Quantitatively, discrepancies between bulk DOM studies and more precise knowledge from SPE-DOM have been noted and remain a subject of debate [*Chen et al.*, 2016]. The carbon extraction recovery is sometimes as low as 10% and varies depending on the specific polarity of each chemical, and as such compositional biases are introduced. Moreover, chemical reactions during the preparation process are not precluded.

Nevertheless, this diverse chemical list illustrates the concept of a chemical continuum where almost every kind of functional group is present within the chemical mixture not as one single compound, but in a succession of small incremental properties [*Lechtenfeld et al.*, 2014; *Noriega-Ortega et al.*, 2019]. This chemical continuum, already huge by the diversity of sources, widens even more as reprocessing occurs. Biological and photochemical processes modify, oxidize, assemble, or disrupt moieties, producing complex pathways of interconnected chemicals which we still barely succeed at modeling even in recent experiments [*Leyva et al.*, 2022].

Overall, the studies on DOM classification have revealed four takeaways. 1) fresh DOM material is energy-dense and is 2) a heteroatomic resource for the food web (Se, P, N, etc.). 3) The mixture of different chemical moieties in the DOM pool explains the observed diversity of physical characteristics (colloidal, jelly, micellar, hydrated, etc.). Finally, the molecular pool diversifies with age as reprocessing decreases compounds molecular size and leads to the formation of ionizable groups (e.g., $R\text{-}COO^-$) [*Cooper et al.*, 2020].

There is a growing body of research on the spatial variance of DOM composition across oceanic transects and ecosystems [*Abdulla et al.*, 2010; *Kowalczuk et al.*, 2013; *Reader et al.*, 2014; *Yamashita et al.*, 2017]. Typically, there is a ubiquitous background of ~40 μM of refractory DOM present worldwide and at all depths of the ocean [*Carlson and Hansell*, 2015]. Yet, another ~30 μM of fresh material is produced at the oceanic surface by phytoplankton and other processes. Some of this biologically available DOM, also termed labile DOM (LDOM), sinks, and provides the substrate for feeding a part of deep-sea life. LDOM is therefore highest at the surface and decreases with depth. This vertical profile is also seasonally driven as the input of labile biochemicals occurs in the North Atlantic after spring and fall plankton blooms. The rest of the time, degradation by-products are the dominant form of DOM in the deep sea [*Reinthaler et al.*, 2013; *Sanders et al.*, 2014; *Zhang et al.*, 2015].

From land to sea, DOC concentrations are highest in rivers and generally decrease with distance from shore. DOM composition differs as well, in that freshwater sources (e.g., terrestrial plants) contain more lignin, tannins, carbohydrates, and amino acids oligomers. Meanwhile, oceanic sources tend to contain more phospholipids and carboxyl-rich alicyclic molecules (CRAM) [*Abdulla et al.*, 2010; *Hansell and Carlson*, 2014]. Beyond the change in biological origin between rivers and oceanic ecosystems, part of this difference comes from abiotic changes like DOM flocculation in the estuarine transition zone. There, the increasing salinity induces the coagulation of amphiphilic exopolymers, with processes like the Ca^{2+} driven DOM self-assembly changing DOM composition [*Abdulla et al.*, 2010; *Verdugo and Santschi*, 2010].

Moreover, mineral surfaces impact much of freshwater DOM composition, while their effect is almost absent in the open ocean. Some minerals have excellent sorption capacities for natural carboxylic acids, such that this sorbed carbon represents a measurable amount of TOC [*Cooper et al.*, 2005]. Hence, clay particles can affect DOM solubility, such as this fraction eventually sediments in lakes' bottoms and estuaries, resulting in its removal from the surface dissolved pool [*Messetta et al.*, 2018; *Vidon et al.*, 2008]. While leaching and removal processes occur, aquatic microbes access mineral surfaces, enabling the enrichment of the DOM pool in heteroatoms (metals, phosphate, trace nutrients, etc) [*She et al.*, 2021]. Hence, even if DOM and soil organic matter diagenesis are processes that should be distinguished, it is interesting to note that stream sediment transport constitutes the process connecting their respective biogeochemistry [*Gueguen and Kowalczuk*, 2014; *Lehmann and Kleber*, 2015].

The DOM composition is also a function of its age and of the conditions in which the reprocessing occurs [*Flerus et al.*, 2012]. Effectively, DOM is in essence a reduced chemical form that can serve as an electron donor for microbial energy production, and a protective substrate against oxidative stress [*Nieto-Cid et al.*, 2005]. Hence, the degradation by-products formed in oxic

conditions tend to be more oxidized than the starting LDOM material. In contrast, in habitats that are anoxic or hypoxic, (e.g., lakes bottom, estuaries' nepheloid layer, etc.), the microbial community uses slower anaerobic processes as they must rely on electron acceptors other than oxygen. As a result, these conditions increase the preservation of fresh DOM compounds [*Hébert and Tremblay*, 2017]. For example, *Jessen et al.* [2017] observed more phosphoester links and less oxidized material in the Black Sea anoxic zone. There were also fewer aminoglycan and lipids, and more sulfur-containing compounds, a set of properties shown to be related to bacterial classes.

The variability of biological activity holds therefore a seasonal DOM signature that can be observed in freshwater systems and in the open ocean [*Heal et al.*, 2020; *Lynch et al.*, 2019]. The other component of this variability is associated with regional ecology [*Coutinho et al.*, 2015]. In contrast to tropical regions, boreal areas of the ocean tend to have fewer cyanobacteria and heterotrophs but more diatoms. Thus, the production of carbohydrates and lipids is high while dissolved organic Nitrogen (DON) and metabolic rates are lower. This situation explains why less secondary metabolite (vitamins and polyketides), less substrate recycling (decayed biochemicals), and less free amino carbohydrates are observed in boreal regions [*Benavides and Voss*, 2015; *Farag et al.*, 2018; *Richardson and Cullen*, 1995; *Ruess and Muller-Navarra*, 2019]. Finally, the DOM pool in boreal aquatic systems varies strongly throughout the year and exhibits less peptide-like but more humic-like fluorescence than in warmer regions [*Yamashita et al.*, 2017].

Overall, space and time (seasons) appear to be key parameters affecting DOM composition. Chemical benchmarking is not trivial, and categories still undergo debate [*Song*, 2017]. Yet, to better ground the concept of DOM composition as dependent on its ecosystem, the case of the northwest Atlantic (NWA) can be used. First of all, a significant part of its DOM is terrigenous in origin (based on $\delta^{13}C$ from NMR spectroscopy) [*Raymond and Spencer*, 2015]. Studies by *Walker et al.* [2009] and *Lechtenfeld et al.* [2015] claimed the presence of many lipids at the surface (45%

of the pool according to both ^{1}H- and ^{13}C-NMR), and carbohydrates (15%) originating from diatoms and Gram-positive bacteria. CRAM forms about 30% of the pool, while unsaturated and polyaromatic compounds are particularly low (less than 5%). Generally, most compounds were found in monomers and oligomers instead of dense polymers. Moreover, in a typical NWA sample: 1 to 5% of DOM mass is labile (LDOM, turnover less than a week); 20 to 30% is semi-labile (less than a year) and 70 % is refractory (RDOM; *Anderson and Amon* [2015]). Those numbers are supposed to vary seasonally as a function of irradiance, biological activity, deep mixing, and riverine inputs. For this thesis, I called those large variations: chemical signatures, sometimes also termed here as imprints. They do not correspond to the presence of specific compounds, but the wider picture of the covarying groups of chemical characteristics in a complex DOM pool.

1.2 Biotic and abiotic reprocessing of DOM and its impact on the carbon cycle

Besides the spatial and seasonal variation of DOM composition, biotic (i.e., the microbial community) and abiotic (i.e., photochemical) processes are also key drivers of the DOM fingerprint, as reviewed in *Carlson and Hansell* [2015].

To start with, plankton releases DOM through active (i.e., direct exudation, and the solubilization of particulate organic carbon (POC)) and incidental processes (i.e., cell lysis, and grazing). Those processes represent 5 to 35% of the primary production per day [*Duursma*, 1963]. Direct exudation causes the release of many compounds by passive diffusion throughout membranes and active excretion. While surrounding water DOC concentration range at best at the micromolar scale, cytosolic concentrations are usually 3 orders of magnitude higher [*Carlson and Hansell*, 2015], and material diffusion following the gradient can occur for low molecular weight molecules [*Bjørrisen*, 1988]. However, a contrasting hypothesis, called the “overflow model” by

Fogg [1966], proposes an exudation of sugars when the primary production exceeds the cell carrying capacity. Starting from the observation that the release of exudates is correlated to the light cycle, and that cytosol and phycosphere composition differ (as seen by the C/N ratio), Fogg proposed an active exudation of carbon-rich compounds occurs when a dissolved inorganic nutrient becomes limited.

Since then, many other factors were brought to the discussion by *Williams* [1990]; *Berman-Frank and Dubinsky* [1999]; and *Mitra et al.* [2014]. Beyond identifying multiple stressors that can induce this overflow of carbon as a means to preserve the producers' physiological functions, they discovered that the process itself fed a wide community of bacteria whose performance was coupled to the success of the phytoplankton itself. From an evolutionary perspective, it is easy to understand how unicellular organisms have favored feeding on the incidental release of cytosolic compounds as this requires little metabolic energy. Yet, the wider picture holds an evolutional benefit for the active releaser, as the release enables a constant exchange of substrates within the planktonic assemblage.

To name only one example of such benefits: nitrogen fixation within cyanobacteria colonies uses LDOM as a conveyer of resources to favor cooperation within the colony [*Seth et al.*, 2021]. Heterocysts develop periodically within longer chains of cyanobacteria vegetative cells [*Kumar et al.*, 2010]. On one side, heterocysts fix more N than they use, but at a higher energetic cost than their anaerobic metabolism can support [*Brown et al.*, 1989]. On the other side, vegetative cells have efficient photosynthesis but lack the capacity of fixing N, which makes them dependent on an external source for their activity. The two kinds of cells engage in a trade-off of reduced carbon in exchange for N-containing substrates (i.e., dissolved free amino acids).

Hence, more than just scavenging by osmotrophs, the exchange of compounds in and out of biomass holds a biological imprint [*Mangal et al.*, 2016]. More generally, active exudation has

also been shown as strategies for microbial biofilm building (polysaccharides), cell communication (auto-inducers), antibiotic weapons (i.e., DMSP), cell removal of the waste metabolic product, and even metal chelation (i.e., siderophore), as a strategy used by microbes for iron absorption [*Hirose*, 2007; *Moran and Durham*, 2019; *Rolland et al.*, 2016].

Another good source of DOM from plankton is cell lysis, where the whole content of cytosol is released by membranes collapsing. The two most common pathways are 1) viral lysis and 2) predation (e.g., by dinoflagellates). As lysis removes 5 to 40% of the ocean's prokaryotic stock each day, it is thought to release 3 to 20 billion tons of DOC per year (worldwide) [*Wilhelm and Suttle*, 1999]. Cytosol provides many crucial chemicals such as DNA, dissolved free amino acid (DFAA), peptides, and monosaccharides as well as colloidal lipids (cell debris), and is therefore of high metabolic value to heterotrophic bacteria.

Last but not least, POC solubilization happens through the effect of extracellular enzymes like peptidoglycan hydrolases, proteases, phosphatases, chitinases, glucosidases, etc. Their collective action increases the DOM diversity and diminishes the average size of its constituents. Enzymes also contribute to a part of the oxidation of fresh material, and the dissolution and recycling of POC content in nitrogen and phosphate (part of at least). Certain studies observe therefore that 50 to 98% of the mass of a sinking particle can be solubilized by this mean [*Smith et al.*, 1992].

Hence, much of the biological exudates are constituted of dissolved free neutral sugars (DFNS), peptides and DFAA, organosulfur and organophosphorus compounds. The evidence shows, however, that those compounds do not accumulate in the water column, as rapid recycling and/or remineralization occurs. For example, the release of extracellular DNA in the ocean is estimated at 13 Pg per year [*Dell'Anno and Danovaro*, 2005]. Yet experiments have shown that its residence time in the water column is only 10 hours, which explains why the *in situ* concentrations

are about 10 μg/L (~3% of the DOC pool) [*Vlassov et al.*, 2007]. This uptake of extracellular DNA is driven by its nutritional value, such as extracellular DNA, which provides 50% of the cell requirement of P despite its minority in the pool [*Dell'Anno and Danovaro*, 2005].

The existence of a labile subset of the DOM (LDOM), which supports cellular competitiveness and maximizes the growth efficiency of the overall microbial community, contrasts with the presence of material refractory to degradation in the DOM pool (RDOM). In fact, an important part of the oceanic pool is refractory, and models suggest its persistence over centuries. This pool is on average much older than oceanic overturning cycles and forms a long-term carbon sequestration reservoir [*Carlson and Hansell*, 2015]. NMR analyses quantified that a third of this RDOM is constituted of CRAM [*Lechtenfeld et al.*, 2014], and MS studies revealed that 85% of CRAM identifiable molecular formulae were heavier (>432 Da) and more saturated (H/C >1) than the rest of the DOM [*Merder et al.*, 2021]. Possibly, CRAM recalcitrance may be due to the complexity that their catabolism represents in contrast to a low energetic benefit.

According to the microbial carbon pump theory (MCP), planktonic exudates are chemical precursors of refractory material [*Jiao et al.*, 2010]. Sometimes, digestive enzymes do not metabolize an entire molecule but target only a specific functional group of it (e.g., an amine). As the process occurs over and over, compounds lose biochemically reactive sites but do not disappear completely. This MCP theory provides a good explanation for the changes in stoichiometric composition from LDOM to RDOM, where the preferential degradation with little investment in carbon induces progressive increases in C/N and C/P with DOM age.

The shift in the chemical composition of DOM has been suggested as a reason for RDOM's refractory nature, yet, catabolic pathways for tertiary carbons, cyclic, and/or aromatic compounds exist [*Gibson and Harwood*, 2002]. Moreover, recent experiments showed that deep-sea DOM, refractory to deep-sea microbiome metabolism, can be quickly metabolized by coastal plankton

[*Shen and Benner*, 2018]. This suggests that the material was not by nature hard to metabolize or low in nutrients and energy, but that deep-sea microbe may not possess the metabolic capability or need to exploit such substrates (i.e., context-dependence). That said, the debate is not settled, and yet another alternative perspective proposed that part of this slower degradation comes from the required remodeling of the bacterial community when adapting to the addition of aged material [*Lennartz and Dittmar*, 2022].

Despite microbial reprocessing being an important part of the carbon cycle, photochemistry is also a key contributor to its transformation. UV rays can induce chemical reactions via mechanisms well documented both in chemistry and ecology [*Mopper et al.*, 2015]. For example, photo-oxidation in surface waters accelerates the remineralization of RDOM either by working directly on the molecule's covalent bonds or through the production of radical oxygen species. Typically, CO_2 and CO emissions are observed under those photochemical processes, but interestingly not always in the same proportions depending on the seasonal component of the DOM [*Reader and Miller*, 2012]. Changing the redox states of certain moieties, like disulfide bonds or C-C covalent bonds, provokes the fragmentation of fresh DOM into smaller compounds [*Kujawinski et al.*, 2004].

By breaking down biologically challenging structures, photo-oxidation can induce a priming effect for further biological degradation [*Kieber et al.*, 1989; *Miller and Moran*, 1997]. Such photo-enhanced biodegradation was shown to remove 2 to 3% of the RDOM content over a year [*Shen and Benner*, 2018]. Yet, this synergy effect is not always observed as photolytic processes also work in competition with microbes' remineralization of LDOM and may reduce the bioavailable carbon [*Reader et al.*, 2014]. Moreover, photo-oxidation can induce the flocculation of terrestrial DOM, generating the loss of metal-chelating moieties [*Helms et al.*, 2013; *Kothawala et al.*, 2014]. Just like biotic processes, these photo-oxidative processes produce an imprint on DOM quality.

Other abiotic processes like the pyrolization of RDOM at hydrothermal vents [*Hawkes et al.*, 2015] and diagenesis of sinking material [*Ogawa and Tanoue*, 2003] also constitute key processes in the carbon cycle, yet are beyond the scope of this thesis. Interestingly, recent developments in the field of DOM research have started to monitor anthropogenic disturbances in the pool composition through land management, and the release of organic compounds and reactants from industry [*Solomon et al.*, 2015; *Zhuang and Yang*, 2018].

DOM is central to the carbon cycle, its quality as a refractory material (carbon sink), its lability (promoting heterotrophy and carbon turnover), coagulating potential (DOM/POM equilibrium), and biological growth efficiency (BGE) frame the kinetics of the biological carbon pump (BCP). The BCP plays an important role in climate regulation by capturing 2 Pg of carbon per year [*De La Rocha and Passow*, 2014]. Moreover, DOM rich in deoxy-sugar polymers, O-methylated polymers, and proteins coming from transparent exopolymers and other sinking particles provide nutritive substrates to suspension-feeders of the meso/bathy and abyssopelagic zones [*Verdugo et al.*, 2004]. Hence, if climate change was to change the kinetics of DOM degradation, the outcomes would be manifest [*Wagner et al.*, 2020].

Overall, this section demonstrated that in addition to spatial and seasonal variation (e.g., riverine discharge); biological (biotic) and photo-processes (abiotic) leave measurable imprints on its composition. However, how those four environmental drivers come together to produce the fingerprints in DOM that we can measure remains a challenging interpretation. As detailed in the next section, a growing body of research is currently bridging that gap to comprehensively model DOM dynamics in the ecosystem. Inspiring myself by their work, I will propose two experimental designs to further explore the integration of biotic and abiotic DOM reprocessing.

1.3 DOM chemical signature as an environmental monitoring tool

From absorbance and fluorescence properties to atomic and functional group contents, different DOM proxies are used in ecology to follow complex phenomena [*Beaupré*, 2015]. Among the list of successes, monitoring water mass movement, biomass abundance, organic matter cycling, etc. have helped a lot in our understanding of the processes unfolding in riverine and marine ecosystems. Yet, the full picture of how environmental factors integrates within the DOM pool remains to be drawn. Little is known as yet on how properties like mass distribution, functional groups, and optical properties covary across the environment. In the first part of this thesis introduction, biotic factors (microbial), as well as abiotic processes (photochemical, seasonal, and water-body type), were pinpointed as key shapers of DOM composition. I want therefore to study the change in DOM signature, defined as a comprehensive view of DOM mass and composition proxies, under a reprocessing experiment. More precisely, the goal is to cross-link the changes within the system using chemometrics.

The first experiment followed the change in the characteristics of Swedish riverine DOM pools over a year; and linked those *in situ* shifts to photolysis and microbial degradation experiments. I hypothesized that the DOM mass spectra behaved as an additive shape of several chemical components, each independently reacting to environmental factors such as regional hydrology, summertime high productivity and/or degradation rates, and winter preservation processes. If this is true, the effects of those environmental conditions can be delineated using a mathematical approach. Chapter 2 tests and documents the views in which carbon reactivity is the driver of DOM composition across time and space in boreal freshwater systems.

The second experiment (Chapter 3) proposes to study the effects of different phytoplankton groups on DOM reprocessing. Estuarine primary producers are known to place an important imprint on coastal DOM, but likely varying depending on microbial assemblages. Hypothesizing

as well that the DOM profiles will change across the growth (time-dimension), I wanted to monitor DOM optical properties and ^{1}H-NMR spectra across fresh planktonic cultures. Yet, before developing further those experimental designs, a review of the available set of techniques studying the chemical signature of DOM is required.

Chemical characterization of DOM is performed following essentially three approaches: 1) DOM atomic and mass composition; 2) detecting the presence of certain functional groups, and 3) interaction with light [*Derrien et al.*, 2019; *Repeta*, 2015]. As each technique holds limitations and assumptions, the co-clustering of several characteristics (signature) is helpful to pinpoint DOM physical properties that have ecological relevance (light absorptivity, nutrient lability, aggregation potential, etc.) [*Farag et al.*, 2018; *Lynch et al.*, 2019].

DOM atomic/mass composition can be quantified through pyrolysis and/or more advanced atomization technics (e.g., ICP-MS). In short, by combusting a sample and employing an appropriate detector, these devices can give the mass of total organic carbon (TOC), and total nitrogen (TN) present in a certain volume of sample (mg/L) [*Cauwet*, 1999]. Knowing the TOC concentration can be quite useful to monitor primary production and remineralization fluxes as well as to evaluate atomic ratios. Different protocols for this have been validated, and the non-purgeable organic carbon (NPOC) method has shown the highest accuracy for the quantification of oceanic DOM. In an NPOC measurement dissolved CO_2 is removed by sample acidification and purging with CO_2-free air before pyrolyzing.

From TOC and TN content, studies have explored the dynamics of the carbon cycle in the environment, determining the rates of accumulation and depletion [*Anderson et al.*, 2015]. The latter appeared to be quite complex with phenomena like phytoplankton blooms, heterotrophic consumption, POM sinking, riverine inputs, UV-degradation, and the range of molecular reactivity, all affecting TOC and TN. In the study by *Keller and Hood* [2011], 70 parameters (such as the

empirical decay rates) were interrelated to affect 55 variables in a series of multiple negative feedback loops reproducing regional ecosystem and biological successions. With iterative calculations of carbon transfer between biomass, DOM, and inorganic reservoirs, their model reproduced most of the real observations in the Chesapeake Bay. That is a baseline of DOC and DON content, experiencing yearly spikes after spring and fall runoff. Their model highlighted the effect of six highly sensitive parameters. In brief, the partitioning of phytoplankton production and its C/N ratio, the bacterial biological growth efficiency (BGE), the bacterial maximum growth rate, the half-saturation constant of LDOM uptake, and the rate of RDOM UV-photooxidation were all found critical.

This study by Keller and Hood, among many others, confirms the central questions of DOM quality to forecast rates of its degradation. While *Redfield* [1958] discovered that the phytoplankton biomass is on average constituted of 6.6 C: 1 N: 0.06 P, the C/N ratio of DOM can be as high as 40, meaning that this carbon pool is too limited in nitrogen to enable a high BGE. The BGE is defined as the ratio of biomass produced over the DOM mass-consumed and quantifies the mass transfer efficiency through a trophic step [*Giorgio and Cole*, 1998]. BGE empirically integrates many notions such as the C/N ratio, the chemical energy density (ED), the energetic cost/benefit of remineralization, etc. Among those chemical constraints, the abundance of lipids increases the system BGE as fatty acids store two times more energy per gram than proteins and carbohydrates [*Stryer et al.*, 2013].

High lipidic content within the DOM signature, estimated by NMR and/or lipidomics, set therefore the ground for an efficient trophic pyramid. For example, the inverted food-web pyramid of the Northwest Atlantic [*Buck et al.*, 1996; *Pershing and Stamieszkin*, 2019], has been associated with the high aliphatic content of its organic matter, enabling one of the highest ratios of bacterial production over primary production in marine environments [*Ducklow*, 1999].

Hence, assessing the TOC/TN and the NMR spectra of our samples will provide insights into their chemical quality. Operationally, NMR technologies provide a quantitative analysis of DOM functional groups spread out in a spectrum of chemical shifts (δ). The study of the relative proportion of those chemical categories is widespread in DOM studies [*Repeta*, 2015]. ^{1}H-NMR can provide quantifiable insights on the aliphatic content of the pool (carbon chains: 0.4-1.6 ppm; CRAM: 1.6-3.0 ppm) to compare its proportion relative to the unsaturated structures (olefinic: 5.2-6.0 ppm; aromatic: 6.0-9.5 ppm) [*Hertkorn et al.*, 2013; *Nebbioso et al.*, 2014]. ^{1}H-NMR also reveals protein-like signals with peptide α-carbons' proton peaks (4.1-5.2 ppm), and amide N-H bonds from 7.8 to 8.4 ppm [*Pisani et al.*, 2015]. Bands from 3.0 to 4.2 ppm integrate the signals from single oxygenated units (methoxy derivatives, alcohols, ether, etc.), and many studies have shown its relation to carbohydrate content in natural DOM [*Lam and Simpson*, 2008]. Further evidence of carbohydrates comes typically from the sugars' free anomeric protons, and related structures, which appear around 4.6 and 5.2 ppm (anomer β and α respectively), or around 5.4 ppm within glycosidic bonds.

Unfortunately, ^{1}H-NMR as 1D spectra remains poorly resolved for the variety of moieties. Yet, further studies have coupled the information of ^{1}H and ^{13}C NMR into HSQC plots to distinguish functional groups [*Hertkorn et al.*, 2013; *Nebbioso et al.*, 2014]. Analytical chemistry developed a variety of approaches like COSY, DOSY, TOCSY, DEPT, etc., yet their use is limited for DOM samples as analytes concentration are typically too low. To maintain the quantification capacity such dilute samples require increasing the number of scans, which in turn limits the number of samples possible to analyze in a timely manner. Thus, we chose to preclude those insightful correlative methods.

Another approach to DOM assesses its optical properties with the absorbance (200 to 800 nm), and the fluorescence spectra of the DOM (UV-vis range as well). The wide absorptivity

of DOM is caused by the diversity in length of conjugated systems, and a subfraction of this operationally defined chromophoric DOM (CDOM) can re-emit the light at a different, lower energy wavelength than the one they absorb (fluorescent DOM; FDOM). This Stokes shift is particularly efficient with aromatic compounds (stabilized excited states and low vibrational freedom [*Valeur and Berberan-Santos*, 2012a]). The core part of the signal is therefore generated by polyphenolic compounds (e.g., lignin residues) and certain amino acids (tyrosine, tryptophan, phenylalanine) [*Murphy et al.*, 2014; *Stubbins et al.*, 2014]. Furthermore, an important diversity of fluorochromes such as the pyrene family, tetrapyrroles, vitamin B, etc. have also been found in DOM studies [*Wünsch et al.*, 2015; *Wünsch et al.*, 2017].

Although CDOM spectra remain an unresolved mixture of information, values at certain wavelengths were mathematically correlated to characteristic concentrations. For example, the CDOM concentration can be approximated using the absorbance at 350 nm [*Mannino et al.*, 2019]. The same can be done with 280 and 664 nm correlating respectively to lignin and Chlorophyll α. Moreover, the specific ultra-violet absorbance at 254 nm ($SUVA_{254}$) value correlated to the proportion of aromatic compounds in the pool [*Weishaar et al.*, 2003]. It was calculated as $SUVA_{254}= \frac{a254}{[DOC]}$ where a_{254} is the absorbance at 254 nm in AU. The relationship they found (positive linear; $R^2= 0.97$) applied from 8 to 38% of aromatic relative to DOC concentration.

UV-vis absorbance also informs the compounds' mass distribution in DOM, through the technic of the spectral-slope ratio (S_r) [*Helms et al.*, 2008]. This parameter is calculated by estimating the slope in 2 regions of the spectra, first between 275 to 295 nm and then between 350 to 400 nm. The ratio of these slopes (usually between 0.8 to 2.5) is correlated ($R^2= 0.86$) to the relative proportion of low and high molecular weight CDOM ($LMW_{<1kDa}$ / $HMW_{>1kDa}$). The theory is that heavier compounds tend to have longer conjugated systems that absorb at long

wavelengths resulting in shallower absorbance spectra. In other words, a high S_r, corresponding to a steep spectral shape, indicates an important proportion of CDOM at less than 1 kDa. The result can be then derived into a percentage of LMW over total CDOM.

The fluorescence properties of the DOM are also insightful [*Coble et al.*, 2014]. Recorded as a 3D excitation-emission matrix, the FDOM signal can be deconvoluted into underlying components using a parallel factor analysis (PARAFAC) [*Murphy et al.*, 2013]. This enables the identification of different regions of the fluorescence signal and allows their associations with particular chemical and ecological properties to be identified [*Valeur and Berberan-Santos*, 2012b]. For example, PARAFAC components correlated to DOM origin and proved successful to distinguish terrestrial DOM from oceanic [*Murphy et al.*, 2008]; and fresh material from older [*Jørgensen et al.*, 2014].

FDOM molecular co-clustering is also an efficient answer to the chemical continuum issue, and it has shown better results than high-resolution techniques. For example, *Amaral et al.* [2016] documented how source changes, DOC and CDOM abundance, FDOM, and aromaticity content were linked to the abundance of *Alpha-*, *Beta-*, *Gamma-proteobacteria*, and *Cytophagaflavobacteria* in a tropical estuary. Observing that those changes match *in vitro* microbial behaviors of oligo/copiotrophy, as well as allochthonous/autochthonous material preference, they conclude that the DOM quality was an excellent proxy of community structure when phylogeny requirements remain low.

It appears therefore that DOM properties can provide insights into the microbial dynamics in the ecosystem, and studies have correlated the plankton community remodeling to DOM properties across seasons [*Coutinho et al.*, 2015; *Heal et al.*, 2020]. That said, concerns regarding fluorescence analyses are sometimes highlighted as it emphasizes the indistinct and controversial molecular family of the humic substances without the power to really identify their diversity [*Ateia et al.*,

2017]. Hence, while our experimental designs are based in a great part on PARAFAC, the use of complementary identifiers is necessary.

Last but not least, we considered using chromatography technologies to simplify the DOM pool and reinforce chemical identification. Chromatography works by chemical affinity and many matrixes exist to resolve several kinds of compounds. The separation can be coarse (molecular size) or extremely precise (up to enantiomeric level) [*Fitznar et al.*, 1999]. The two main separation methods are gas chromatography (GC) and liquid chromatography (LC). While GC is very efficient for lipids, LC is more universal [*Kennicutt and Jeffrey*, 1981].

The choice of the matrix and the eluant is mostly impacted by the assumptions made about the nature of the analyte, such as the separation will only be optimized for a subset of all the analytes in DOM. While GC is appropriate for volatile compounds, an important part of DOM would stay unassessed. The most generalist approach is therefore to use size exclusion chromatography (SEC) [*Hawkes et al.*, 2019]. Yet, the resolution remains low as the thousands of compounds from a chemical continuum are eluting in just a few minutes. High-pressure conditions increase the separation a bit, but so far, no gold technique has been internationally recognized, and it is common to see a diversity of approaches in the literature (i.e., reverse-phase LC, asymmetric flow field flow fractionation, etc.) [*Sandron et al.*, 2015].

Overall, chromatography is useful to simplify the pool but implies many analytical assumptions as the analytes' solubility and behavior with the mobile phase. Such assumptions are not always verified, yet it helps in separating DOM prior to its analysis under mass spectrometry. An LC-MS like this detects the m/z value (mass-to-charge ratio) and the abundance of ionized compounds. Ionization methods are diverse, but the ESI approach is preferred, as this so-called soft ionization limits molecular fragmentation. For my research, I had access to an SEC-HPLC-ESI-MS dataset that we aimed to further resolve using PARAFAC.

Overall, the chemical parameters we gather about DOM look for categories of absorbance (a_{350}; $SUVA_{254}$; S_r; spectra shape analysis), fluorescence spectra, atomic content (TOC; TN), and chemical functional groups (NMR), and affinities for chemical surrounding (LC; solid-phase extraction). Furthermore, multivariate data-analysis approaches like PARAFAC, helped us to cross-link the information between analyzers and environmental conditions. Among those, biological oxygen demand (BOD), inorganic nutrients, microbial content, and riverine discharge were identified as key elements to monitor as they have important impacts on DOM quality [*Abdulla et al.*, 2010; *Hansell and Carlson*, 2014; *Solomon et al.*, 2015].

1.4 Experimental designs

This thesis aims to document through two separate experiments how the DOM properties change across a range of biotic and abiotic factors. Knowing that each condition will hold a unique set of optical properties, mass-spectra, and ^{1}H-NMR spectra, I hypothesized that their signature would be the summation of the environmental factors they have been exposed to. The first experiment explores the extent of such assertion while the second detail the effects of more phytoplankton production and related biological processes. In contrast to approaches seeking specific compounds to biomark the ecosystem (e.g., *Taipale et al.* [2016]; or *Zhang et al.* [2017]), I aimed to analyze a more comprehensive set of properties.

In brief, here are the abstracts of these experimental designs:

Chapter 2: Spatial/temporal variability in carbon reactivity of Swedish rivers

I used an unpublished dataset collected in 2012-2013 by Heather E. Reader and her collaborators to distinguish the spatial/temporal signature of DOM. The DOM of three Swedish rivers was monitored monthly for a year, and extra volume was taken for photolysis and microbial

degradation studies. The dataset totaled 100 samples that were analyzed in an SEC set-up in line with three detectors taking 1) the absorbance, 2) the fluorescence, and 3) the mass spectrum.

After data QA/QC pre-corrections and outlier removal, I normalized and associated the data from the three distinct detectors into one tridimensional matrix of 85 samples by 190 time-points (i.e., retention time) by 2185 variables (either a mass in Da or an absorbance/fluorescence wavelength). From there and inspiring myself from the work of *Reader et al.* [2015]; *Solomon et al.* [2015]; *Lynch et al.* [2019]; and *Wünsch and Hawkes* [2020], I modeled the data using a PARAFAC method.

The results suggest the presence of four distinct DOM signatures whose characteristics were making sense physically in the ecosystem. The relative abundance of those components varied seasonally and across the three rivers. Such variations were coherent and could be explained in light of the degradation experiments. The quantitative changes in composition were supported statistically by an ANOVA, and we found correlations of sample score to seasonal riverine discharge and BOD. Variable loadings were used to distinguish the 4 fingerprints in the mass domain (mass region, and ion family periodicities), the absorbance (relative absorbance, and S_r), and fluorescence (relative emission intensity and wavelength).

Overall, we found distinct reprocessing dynamics within those riverine DOM sub-pools, and each was associated with a mass and optical signature. Linear combinations of those properties explained up to 96% of the data variance, which supports the hypothesis of true chemical components in the environment, each composed of thousands of interrelated chemicals, and subjected to degradation factors with varying sensitivity.

Chapter 3: DOM exudates' properties across the growth of coastal diatoms

In that second project, we aimed to document how the coastal DOM of Newfoundland may change under the forecasted size-diminution of the planktonic community. Due to climate change,

and its effects on coastal nutrient inputs, diatoms species proportion and balance with smaller phototrophs will change [*Balaguru et al.*, 2018; *Buck et al.*, 1996; *Daniels et al.*, 2015; *Greene et al.*, 2008]. Hence, imagining an experimental design like the work of *Amaral et al.* [2016]; *Guruprasad and Ramachandra* [2011]; *Heal et al.* [2020]; and *Lechtenfeld et al.* [2015], the coastal water was fractioned into three size groups of plankton to compare their exudates.

All groups were produced in independent triplicates, for a total of 9 bottles, and we encouraged planktonic growth by spiking with inorganic nutrients up to the normal spring concentrations in the region. Growth in that controlled environment (light and temperature) was followed every 2 days by *in vivo* chlorophyll fluorescence. Across the experiment, we took 4 DOM profiles (a multi-proxy signature), corresponding to the experiment's start; the early growth, the end of growth, and the plateau/decline stage. Each DOM profile assessed the optical properties (absorbance, fluorescence), DOM elemental composition (TOC, TN, inorganic nutrients, and chlorophyll α), and its ^{1}H-NMR spectrum.

In contrast to our expectations, the plankton size fractionation failed to separate plankton taxa. Instead, similar communities of chain-forming centric diatoms developed in all treatments, so we recentered our data analysis around DOM trends across the growth. Essentially, the bulk of DOC exudation was carbohydrates-related and released in late bloom stages. Moreover, five fluorescent components were identified and produced in contrasting patterns: some parallel to DOC exudation and others earlier in the growth phase.

Among all, we found differing behavior of the two proteinaceous-related FDOM components, with C_{310} disappearing from the water column quicker than C_{335}. Moreover, $SUVA_{254}$ appeared to behave as a function of the relative activity of heterotrophs and diatoms (which released the majority of the DOC). Altogether, those elements highlighted the relationships

between planktonic processes and DOM characteristics, such as its *in-situ* chemical quality could be highly sensitive to the N/Si ratio.

◊◊ Concluding remark ◊◊

The two following chapters of this thesis are manuscripts for publication, followed by an overall discussion and the perspective of those findings. While both experiments investigate DOM signatures through quite different contexts, they form a complementary set of designs that investigate first the abiotic/biotic/refractory processes and then develop the understanding of biotic ones using several microbial assemblages. They also allow the comparison of freshwater and coastal environments, *in situ* characteristics and *in vitro* reprocessing; and bulk-DOM techniques against *in silico* chemometrics of SPE-DOM. Covering this wide set of conditions enabled us to obtain a large understanding of how biotic and abiotic factors affect DOM characteristics of mass, absorbance, fluorescence, and functional groups.

Although a lot remains to explore, it seems clear at the end of this work that the DOM intrinsic properties are less erratic than they appear. Hence, further biogeochemical models should better account for the variation in those properties under a changing climate.

Chapter 2: A PARAFAC approach to dissolved organic matter characterization reveals linkages between chemistry and reactivity in Swedish rivers

Authors and affiliations *Manuscript*

Antoine V. Viallard [1]; Nikoline J. Nielsen [2]; Colin A. Stedmon [3]; Emma S. Kritzberg [4]; and Heather E. Reader [5]

[1] Environmental Science Program, Memorial University of Newfoundland, St John's (NL), Canada

[2] Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark

[3] National Institute of Aquatic Resources, Technical University of Denmark, Kgs Lyngby, Denmark

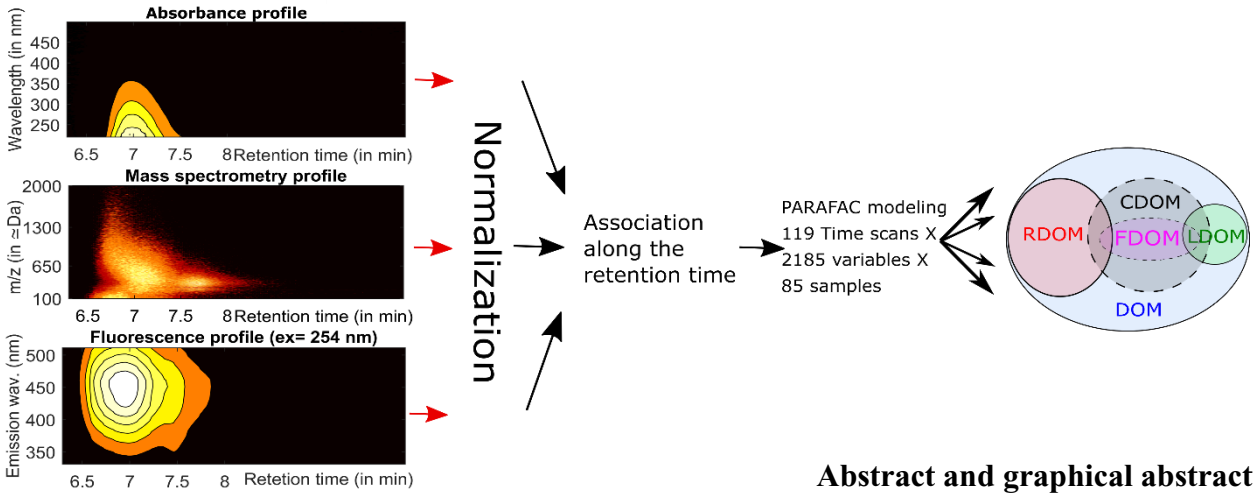
[4] Department of Biology/Aquatic Ecology, Lund University, Lund, Sweden

[5] Department of Chemistry, Memorial University of Newfoundland, St John's (NL), Canada

2.1 Co-authorship statement and acknowledgment:

Author contributions: HR, CS, EK conceived and designed the experiment. HR performed the fieldwork, experimental work, and sample preparation. NN ran the UPLC system and made the preliminary data corrections. AVV performed the computation and the data analysis, with the input of all authors on data interpretation. AVV wrote the manuscript with significant input from the other authors. This manuscript is provided with the approval of all intellectual property owners.

Acknowledgments: The field experiment was financed by the Swedish Research Council (VR, no.2010-4081; granted to EK), the Managing Multiple Stressors in the Baltic Sea project, FORMAS grant no.217-2010-1267, and the Danish Research Council for Independent Research (DFF 1323-00336). The computational work was supported by the Natural Science and Engineering Research Council of Canada (CGS M-2021 to AVV and the Discovery Grant and Canada Research Chair Programs to HR).



Abstract and graphical abstract

The chemical complexity of dissolved organic matter (DOM) hinders progress in understanding its role in the global carbon cycle. Properties derived from a range of techniques such as mass spectrometry and molecular spectroscopy are usually studied in separated boxes, challenging our capacity to link chemistry and ecology. Computational advances, such as PCA, help to identify those hidden links but are limited to bidimensional datasets. By contrast, here we used PARAFAC to study the DOM composition with a multidimensional detection strategy.

Riverine DOM from three Swedish rivers was collected monthly for a year to form a comprehensive dataset covering different climatic zones, seasons, and exposure to both biotic and abiotic degradation. Using size-exclusion chromatography (SEC) with three detectors (absorbance, fluorescence, and mass spectrometer) we assessed DOM composition. The time-resolved data were normalized and associated along the retention time, then a PARAFAC analysis deconvoluted 4 chemical fractions which demonstrated relationships to sources and reactivity in the environment.

A figure assembling a range of spectral data was constructed and provided answers on key DOM sub-pools as the “chromophoric” DOM, “fluorescent DOM”, “refractory DOM” and “labile DOM”. While these sub-pools displayed spatial and seasonal DOM variance which was consistent with their chemical reactivity under degradation, the parsimony of our solution suggests that compounds' proportion may be closely interrelated by still unknown mechanisms.

2.2 Introduction

Dissolved organic matter (DOM) in freshwater systems constitutes the complex substrate used by microbes at the base of the food web. It is released by many processes: from soil organic matter leaching to autochthonous phototrophic activity. Operationally defined by filtration of water, DOM encompasses many thousands of unique compounds of which chemical properties are still largely undifferentiated. The downstream transport of DOM to estuaries forms a key part of the carbon biogeochemical cycle [*del Giorgio and Pace*, 2008; *Hansell and Carlson*, 2014].

Typical chemical proxies used to characterize the DOM range from assessing bulk properties – such as absorbance and fluorescence spectra – to high-specificity chemical methods such as nuclear magnetic resonance (NMR) or mass spectrometry (MS) [*Abdulla et al.*, 2010; *Hertkorn et al.*, 2013; *Lechtenfeld et al.*, 2015]. The development of high-resolution chemical methods for DOM characterization has made progress in identifying molecular properties and clusters. Yet, as these techniques rely on solid-phase extraction prior to analysis (most of the time), they inherently constrain the extent of interpretation to the extractable fraction [*Hawkes et al.*, 2020]. As a result, the operationally defined subgroups known as "chromophoric" DOM (CDOM), "fluorescent DOM" (FDOM); high molecular weight compounds (HMW-DOM), etc. are typically studied in separate boxes [*Derrien et al*., 2019]. This creates discrepancies in information, leaving important uncertainties in linking ecosystem-based interpretation between different studies. For example, studies like *Kothawala et al.* [2014] can screen the DOM quality on large geographical and temporal scales using low-cost measurements to identify the dynamics of broad DOM properties. On the other hand, *Leyva et al.* [2022] used recently an *in-silico* model to show the possible synthesis/degradation pathways between 9,000 compounds from a single DOM sample given the resource-intensive technique of MS-fragmentation. Yet, assembling the information of both studies would be speculative as little relation between mass and optical spectra is known.

DOM degradation is linked to both biotic and abiotic factors [*Carlson and Hansell*, 2015; *Cory and Kling*, 2018]. In surface waters, photochemical processes can turn over significant amounts of DOM to inorganic carbon (i.e., CO and CO_2) [*Reader and Miller*, 2012]. Meanwhile, microbial reprocessing preferentially affects nitrogen, sulfur, and phosphorus-rich compounds, as well as very light (<300 Da) and heavy (> 1kDa) compounds [*Amon and Benner*, 1996]. Despite this broad understanding, factors dictating individual substrates' reactivity or lack thereof remain elusive as studies show remarkable metabolic plasticity and context-dependence of degradation processes [*Gibson and Harwood*, 2002; *Jiao et al.*, 2010; *Shen and Benner*, 2018].

Chemical transformations of DOM are not necessarily in steady-states and are affected by daily/seasonal variations, as well as differences in local nutrient sources and catchment hydrology. Those variables are patchy and can affect DOM characteristics over temporal and spatial scales [*Zhuang and Yang*, 2018]. DOM is therefore a dynamic mixture that is constantly remodeled by environmental factors, creating a continuum of structural moieties.

In response to the complexity of the DOM pool, chemometric approaches have been proven useful in identifying covarying variables as well as those with explanatory potential. Parallel Factor analysis (PARAFAC) has been widely applied to FDOM spectra and has furthered our insight into DOM biogeochemistry [*Murphy et al.*, 2013]. PARAFAC works by distinguishing mathematically non-covarying proportions of components [*Bro*, 1997; *Murphy et al.*, 2018]; and has been suggested as a promising approach for interpreting HPLC-MS data.

Due to the chemical continuum nature of DOM, chromatographic experiments usually produce poorly resolved separations, but the variable clustering capacity of PARAFAC should theoretically reveal compositional interactions between chemicals' properties [*Bro et al.*, 2010; *Bylund et al.*, 2002; *Dąbrowski*, 2020]. However, this exciting possibility of reducing tridimensional dataset complexity remains widely underexplored, with *Wünsch and Hawkes* [2020]

pioneering the method in DOM studies. As an example of the practical knowledge these models yield, the monitoring of the transport of different metal ions in an aquatic system revealed a direct relation to mass fingerprints (Van Krevelen), where the proportion of modeled components across the elution was coupled to change in polarity and chelation ability of DOM samples.

In our case, we proposed to investigate PARAFAC potential within a complex environmental design in which the DOM pool of three Swedish rivers has been monitored monthly for a year coupled with degradation experiments. The systematicity of this design has proven useful in detecting Spatio-temporal changes in DOM mass spectrum [*Reader et al.*, 2015], and the variety of environmental factors covered, such as catchment size and climatological characteristics allow a fair view of the extent and regulation of environmental processes across the Swedish freshwater system.

However, we propose this time to push further the analytical limit using size-exclusion chromatography (SEC-HPLC) coupled with three in-line detectors: absorbance, fluorescence, and mass spectrometer. The addition of two variables (chromatography and fluorescence), absent from the original analysis, enabled us to dive deeper into the environmental links driving the chemical composition of DOM. The latter was hypothesized to behave as additive sets of independent sub-pools whose abundance across season and space is linked to their carbon stability, their optical properties, and the hydrological context at the time of sampling. Therefore, a signal decomposition under a trilinear PARAFAC model should be able to distinguish those subgroups of chemical properties.

2.3 Material and methods

2.3.1 Sampling and data acquisition

The experimental design is the same as described by *Reader et al*. [2015]. In brief, the riverine DOM of three Swedish rivers has been sampled monthly from March 2012 to February 2013. Namely, Ume, Emån, and Lyckeby rivers form a diverse group of catchment characteristics in terms of flow volume, climate, and complexity [*Reader et al.*, 2014]. For each, one sample was taken at mid-month, and filtered first with GF/F and then under 0.2 μm with polycarbonate filters (Millipore) on the same day as sampling. Samples were adjusted to pH = 2 with HCl and stored at 4°C until solid-phase extraction on PPL cartridges (Agilent, 2.5 mg of carbon per cartridge).

In addition to the *in-situ* samples, extra volume was taken for degradation experiments. *Photodegradation set-up*: 1 L of sample (ambient pH) was irradiated at 15°C under UV-A centered lamps for 6 days. The assessed months were May, August, October, and January for each of the 3 rivers. *Microbial reprocessing set-up*: Samples filtered only through GF/F (0.7 μm) were placed in the dark at ambient temperature (mean 20.6°C) for 4 months. This procedure kept microbes in, and samples were then spiked with an excess of inorganic nutrients to a final ratio of 45C: 9N: 1P to ensure carbon limitation (*Goldman et al.* [1987]). Microbial reprocessing was conducted on samples from March, May, August, and January.

The SPE extraction for these samples was done using the same volume as for their corresponding *in situ* samples. SPE-DOM was eluted in methanol, dried entirely, and redissolved into a water-based matrix for SEC. The SEC was in line with 3 detectors: a diode-array detector (DAD, 220–499 nm, 1 nm increment, scan rate= 5 s^{-1}); a fluorescence detector (FLR, λ_{ex}=254 nm - λ_{em}=331–511 nm, 10 nm increment, scan rate= 1 s^{-1}); and finally an Ultima Global quadrupole time-of-flight (QTOF) mass spectrometer (Waters Micromass, from 100 to 2000 m/z – 1 m/z resolution, scan rate= 0.87 s^{-1}).

2.3.2 Dataset characteristics and pre-treatment

The dataset consists of 60 samples: 3 rivers × (12 monthly samples *in situ*, 4 others subjected to photodegradation, and 4 subjected to long-term microbial degradation). 40 analytical controls were also added to assess mass accuracy, contaminants from the mobile phase and the solid phase extraction, extraction consistency, and the stability of the preservation method.

The data were imported into Matlab and the data from the three detectors were aligned and synchronized in the retention time (RT) dimension. Delays between detectors were determined using 3 reference points in the chromatograms, and the data were shifted along RT accordingly. Then, the closest DAD and FLR scans for each MS reference point were pinned and the rest of the scans were discarded (i.e., considering a DAD scan at T=599.95 s and another at T=600.15 s, the 1^{st} is used to match the MS scan at T=600.0 s while the 2^{nd} is discarded). 1016 time points remained covering 19.5 min of elution, from which we sub-selected the 6.3 to 10 min time frame to continue our study, corresponding to the entire elution profile of the samples.

Next, the detector backgrounds and off-sets were suppressed. In MS, ion counts lower than 17 were considered background and replaced by 0. The mass spectra for milli-Q water extractions consistently showed 14 m/z values with a signal above the threshold indicating potential contamination from the PPL cartridges used in the SPE. These m/z values were discarded for the whole dataset. For absorbance and fluorescence, instances of detector saturation and strong baseline deviations were noted (i.e., the signal did not return to baseline after elution). This handful of samples was converted to 'NaN' values over the problematic regions. Three tridimensional matrixes remained after this QA/QC, containing the absorbance, fluorescence, and mass spectrum over ~4 min in elution time.

2.3.3 Dataset normalization, association, and PARAFAC modeling

The resulting datasets (RT × detectors × sample) were normalized by mean scaling each variable separately for each detector. Normalizing factors were calculated by averaging signal intensity excluding 0 and NaN values within dataset slices of 1 variable by all retention time by all samples. This method compensates for the differences between detectors' signal intensity while keeping the relative DOM concentration differences between samples. Finally, the detectors were concatenated along the retention time axis to form a single matrix of 100 samples, 119 times scans, and 2185 variables (either a fluorescence emission wavelength, a mass/charge ratio, or an absorbance wavelength).

PARAFAC analysis was conducted from 2 to 6 components. The model was constrained by nonnegativity and a convergence criterion of 10^{-6} was used. The model was validated using residual analysis, split-half analysis, and random initialization as described by *Murphy et al*. [2013]. Tests indicated that 2 samples (Emån and Lyckeby both December - *in situ*) and 13 controls were outliers and should be excluded to model adequately the data. The best model had 4 components and was selected for interpretation.

2.3.4 Deconvoluting the model properties

Each component has three dimensions: the retention time loadings, the samples' scores, and the variables' loadings. The retention time loadings represent the chromatogram of a given component, and it should optimally look like a gaussian peak. Sample scores represent the quantitative proportion of component C_x in each sample. Sample scores were therefore sorted by rivers (Ume, Emån, and Lyckeby), sampling month, and treatments (that is: no treatment, photodegradation, or biodegradation).

The detector loadings, representing components' optical and mass properties, are first separated into the three independent matrixes for m/z ratio, absorbance, and fluorescence, and then unnormalized by multiplying loadings by the normalization factor used initially.

The effects of the degradation experiment were followed by the normalized variation of scores between the *in-situ* sample and the degraded sample for the same river and month. Moreover, the contribution of each component to the variable (y) as a proportion of the modeled signal ($C_x C_{V_y}$) was calculated using equation 1.

$$C_x C_{V_y} = \frac{V^*_{L_y C_x} \times \bar{X}_{SC_x} \times \int T_{C_x}}{\sum_{i=1}^{4} V^*_{L_y C_i} \times \bar{X}_{SC_i} \times \int T_{C_i}} \quad (1)$$

Where $V^*_{L_y C_x}$ is the denormalized variables loadings of component C_x. $\int T_{C_x}$ is its integrated retention time's loadings along with C_x elution profile, and $\bar{X}_{SC_x}$ the averaged sample's scores of C_x. The product of those factors quantifies C_x modeled signal accounting for its relative sample's abundance and time contribution to the DOM pool. This value is unique to each absorbance/fluorescence wavelength or m/z – and is then reported as a proportion to the sum of the partial signal from C_1 to C_4 ($\sum_{i=1}^{4}$).

2.3.5 Comparison of components' properties alongside spatio-temporal variability

To test statistically the effect of spatial/temporal and reactivity variability, sample scores were transferred to an R environment and studied using MANOVA [*Denis*, 2020]. Assumptions were tested using a Shapiro-Wilks normality test of the residuals, and the Bartlett test (homoscedasticity of groups). The non-parametric test of Kruskal-Wallis was used when necessary. Post-hoc analysis was conducted using Tukey's tests (parametric) or Wilcoxon's (non-parametric). Pre-tests used a high standard of rejection (0.1), while statistical significance was considered when $p_{value} < 0.05$ and using the *Holm* [1979] correction method.

The correlation of component 4 scores with river discharge was investigated using daily mean discharge obtained from the Swedish Meteorological and Hydrological Institute's Vattenweb (SMHI, http://vattenwebb.smhi.se/station/). Discharge at sampling days was normalized to the rivers' annual geometric mean discharge. Further on, we reused the biological oxygen demand (BOD) data reported by *Reader et al.* [2014]. These values were obtained using a Mettler Toledo titrator after a two-week incubation experiment (Winkler method).

2.4 Results

2.4.1 PARAFAC model stability

A 4-component model was found to explain 96.5% of the dataset variance with a core consistency of 55%. As expected, most components have a pseudo-gaussian shape along the retention time dimension, in coherence with chromatographic elution behaviors. Quality controls of the model can be found in appendix A. Sample scores were independent in all but one instance (C1 vs C4; R^2 = 0.45), and most of the groups of conditions were normally distributed and homoscedastic, and with effective co-clustering of analytical controls.

Those validation elements support the quantitative stability of components' shapes (Fig. 2.1), and we represented the contribution of each component to the averaged sample signal (Fig. 2.2).

2.4.2 Components' characteristics

Component 1: C1 is mainly composed of m/z ratios from 470 to 1040 Da. While those represent no more than 15-25% of the total mass spectrum signal, it accounts for 70-80% of the absorbance signal. This high absorbance also represents a proportionally low amount of fluorescence (10% of the signal at F_{max}= 450 nm). As C1 is characterized by a strong absorbance signal, it can be classed as a significant subset of the CDOM pool. Noteworthily, all rivers had similar scores, with a depletion during late summer/early fall (p_{value}< 0.01).

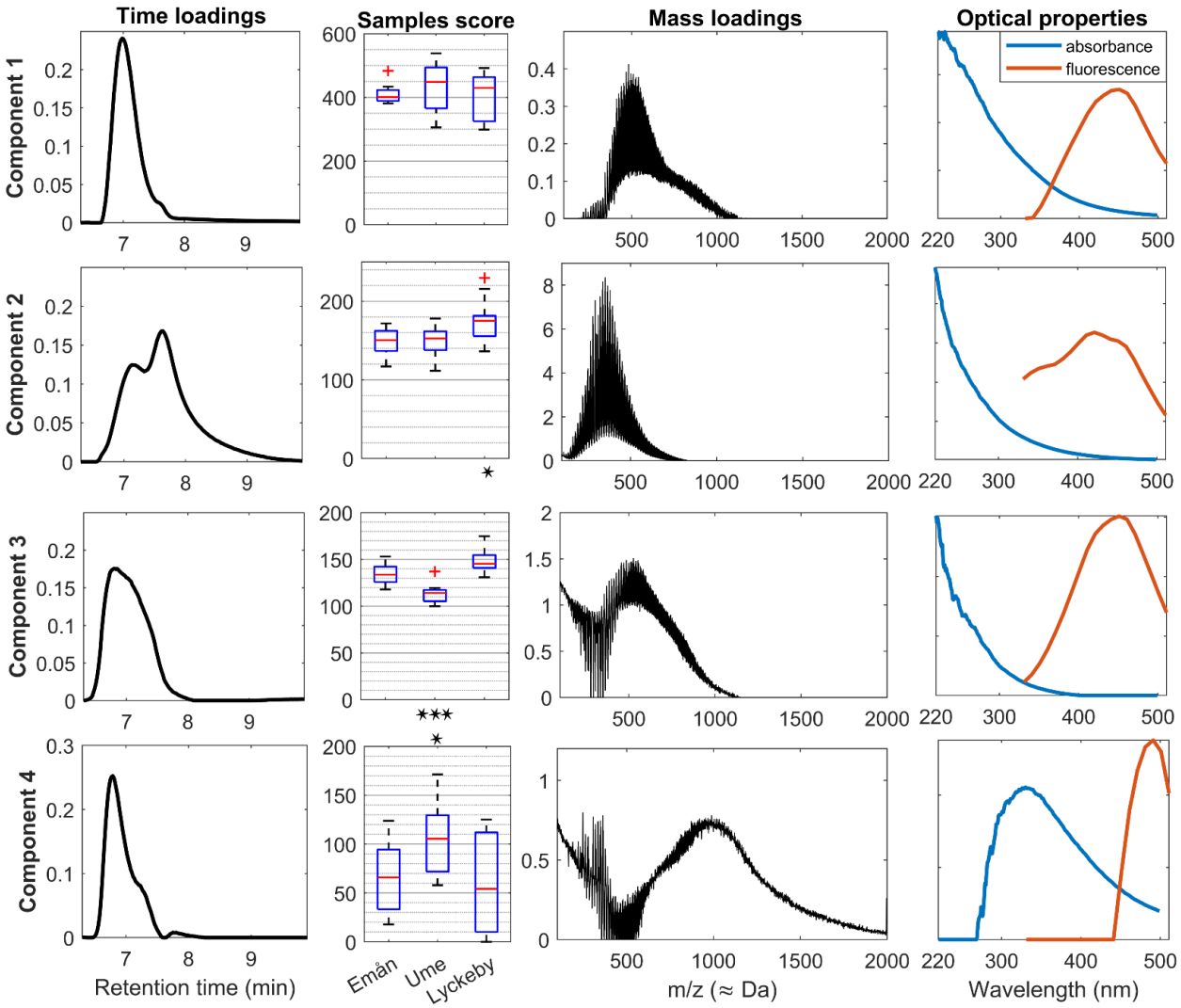


Figure 2.1: Experimental fingerprints for components one to four (vertically). For each – left to right – are shown the elution profile, the sample's score distribution per river across a year, the MS loadings (denormalized), and absorbance/ fluorescence spectra. These latter are scaled to fit together, and spectral slope ratios (Sr) of absorbance shapes were calculated. From C1 to C3, S_r are 0.866; 0.989, and 0.645. MANOVA results of samples' scores differences between rivers are indicated underneath boxplots when significant. Symbol (*) corresponds to $p_{value}<0.05$; and (***) stands for $p_{value}<0.001$.

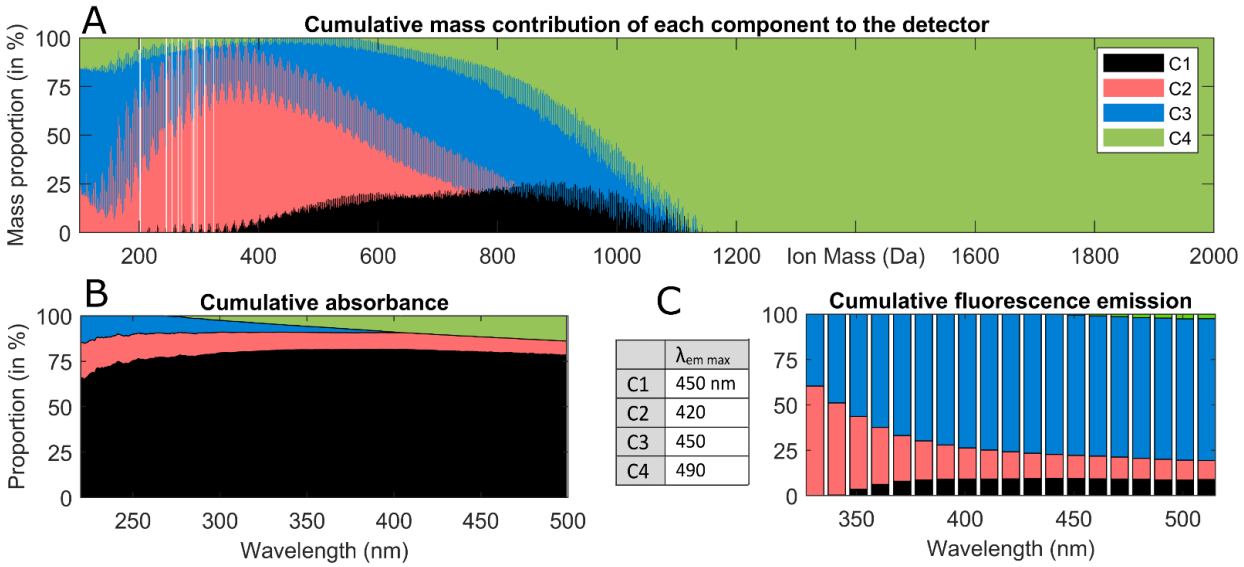


Figure 2.2: Mass and optical properties distributed between the 4 components. In each graph, components' contributions to the detector's signal were calculated and stacked vertically for A) mass values; B) absorbance signal; and C) fluorescence. Values stack up to 100% representing the total explained signal meaning that, for example, Comp. 1 contributes the most to the absorbance signal. In panel C, a table list the emission wavelength of FDOM spectrum maxima for λ_{ex}= 254 nm.

Component 2: C2 forms the core part of m/z from 200 to 600 Da and contributes strongly to the fluorescence at low emission wavelengths. By contrast, C2 accounts for a constant 18% of the absorbance signal across the entire spectrum. It has a bimodal elution shape with a delay of approximately 1 min between peaks. The existence of compounds of the same mass with different affinities to the column is typical of isomer behaviour. Additionally, both peaks are delayed to other components, which indicates higher column interactions of the compounds in C2 compared to the other sub-pools. Finally, the seasonal variation of C2 for all rivers is low, but Lyckeby river was enriched in it (p_{val}= 0.053 after Holm correction).

Component 3: C3 is a subset of compounds abundant between m/z 100 – 280 Da and then between m/z 450 – 1000 Da. It is not a discontinuous bimodality, however, certain ion families – (Table 2.1; appendix A) – have a null score indicating some ion families are strictly excluded. This bimodality of masses may explain the non-gaussian elution shape of C3 in an SEC column.

Importantly, C3 accounts at least for 75% of the total fluorescence signal while playing a modest role in absorbance (17% at low wavelengths and then decreasing until 380 nm). Hence, C3 must contain an essential part of the so-called "FDOM" compounds. Over the entire year, C3 was significantly depleted in Ume when compared to the other 2 rivers ($p_{value} < 10^{-5}$).

Component 4: C4 forms the core of heavy masses (800 to 2000 Da) of the studied DOM pool with 1 non-periodic peak at 248 Da (Appendix B). C4 absorbance represents 17% of the signal at 500 nm, while its fluorescence remains extremely low compared to the other components even at high emission wavelengths. Considering C4's absorbance peak at 340 nm, the excitation wavelength of the detector of 254 nm is likely too low to efficiently excite C4's fluorophores.

2.4.3 Reactivity patterns

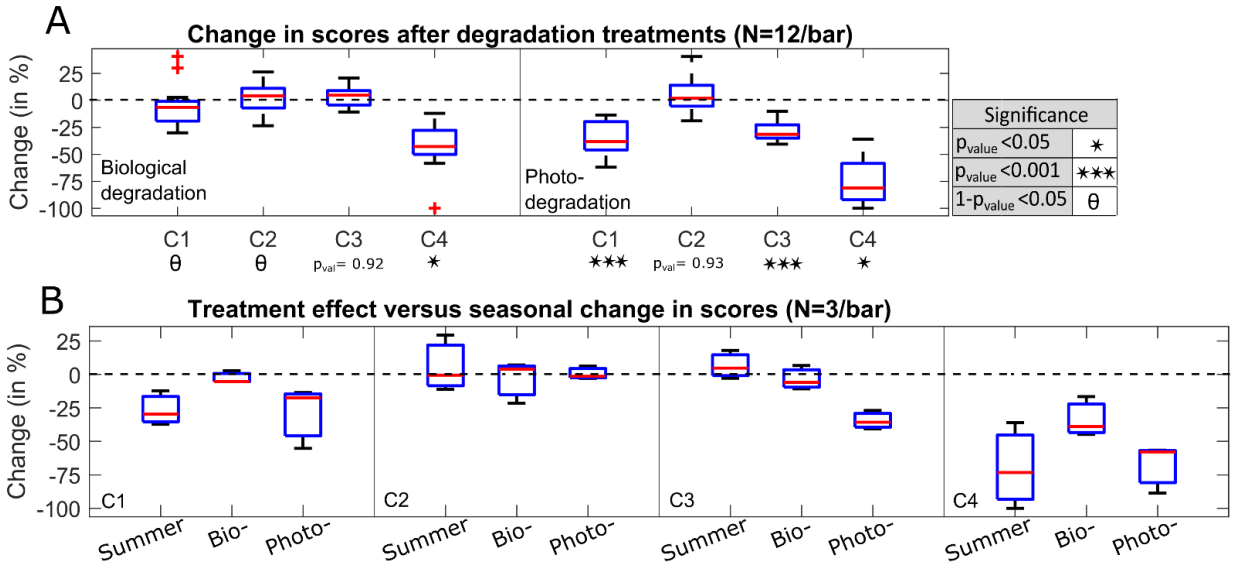


Figure 2.3: Treatments effects compared to seasonal variance. Between each condition, the amount of change in component score was calculated. The results are represented by boxplots within the domain going from -100% (total depletion) to the positive realm (increase), and the dotted line at 0 represents no change in scores. The upper panel displays the results between *in situ* samples and the treated ones, while the bottom panel compares those changes in January treated samples to the "August to January" change. For panel A, the statistical difference significance (p_{value}) and confident similarity ($1-p_{value}$) are symbolized underneath.

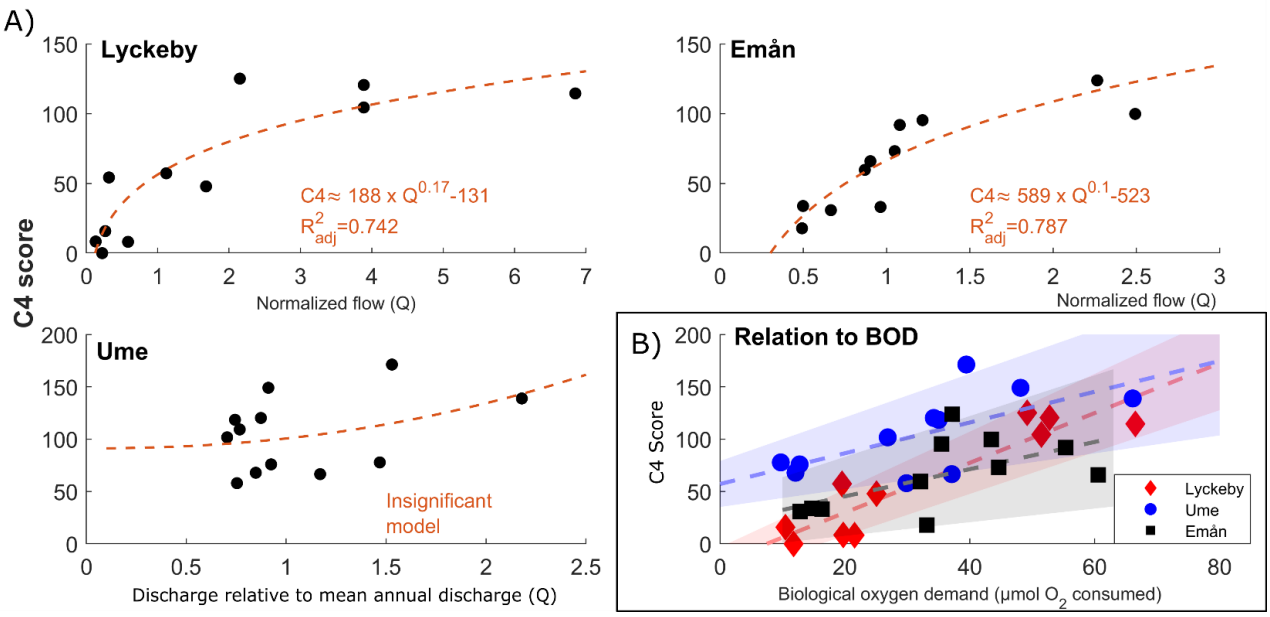


Figure 2.4: Samples score of C4 *in situ* explained by riverine discharge and BOD. A) Regression of C4 score to relative flow volume for each of the 3 rivers (N=11 to 12 per river) was hypothesized to follow an n^{th} root function. The relative flow corresponds to the given flow at the sampling date (in m^3/s) divided by the geometric mean flow of that river. When significant, the model parameters were indicated. In B) C4 was linearly correlated to the samples' biological oxygen demand (BOD). All individual models were significant ($p_{value} \leq 0.075$), and the shaded area corresponds to zones between ± 1 standard error of models' parameters. Overall, a relationship of $C4 \approx 1.7 \times BOD + 20$ ($R_{adj}^2 = 0.406$; $p_{value} = 6 \times 10^{-5}$) was found when pooling the three rivers' data.

Studying the effect of degradation treatments (Fig. 2.3A), only the scores of C4 are affected by microbial reprocessing, and all components but C2 can be photodegraded. The sensitivity to photodegradation is consistent with the previously determined high UV absorbance of C1 and C3. We did not observe any components significantly produced by degradation treatments, but instances of increase exist among certain samples (revealed by a positive variation rate).

Correlation analysis of C4's scores indicates a strong link between its presence and the system hydrology (Fig. 2.4A). In the 2 smallest rivers (Emån and Lyckeby) a significant positive relationship with the riverine discharge could explain ~74% of C4 variance, but no relationship

was observed in the case of the Ume River. A positive linear correlation was also found between C4 abundance and BOD in all rivers, including this time the Ume River (p_{value}= 0.04).

The characteristics of Ume catchment contrast with the other catchments: it is large (roughly 27 000 km^2) and composed of two sub-catchments, one forested and the other open land, while Emån and Lyckeby are relatively small and more densely forested. Ume is also the northernmost river and experiences a sub-zero average winter temperature across the catchment. As a result, this river experience two discharge peaks, the spring freshet in May and a rain-associated event in October, while the other systems experience their main annual discharge peak in the winter months.

Fig. 2.3B shows the relationship between seasons for all components. The differences in C1 from January to August are similar to those obtained by photobleaching January samples. In contrast, C2 scores do not change between winter and summer, which is consistent with its chemical stability. Unlike C1, C3 scores were higher in summer, following an opposite trend to the expected one suggested by its photolability. Regarding C4, both degradation treatments lowered the scores from January (winter) samples but not as much as observed in the seasonal difference (August – January). This reflects C4 complex behavior which also follows riverine discharge as well as environmental degradation factors.

2.5 Discussion

2.5.1 Components' behavior within the biogeochemical context

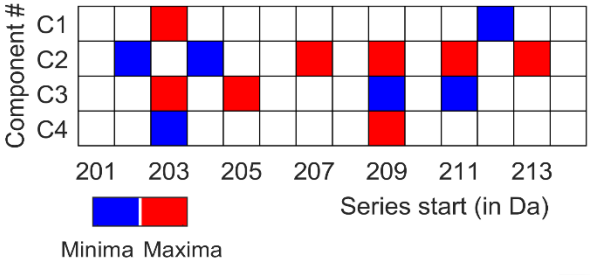


Table 2.1: Fingerprints patterns in ionic masses. We hypothesized that the distributions of mass loadings minima and maxima – observed from 200 Da to 450 Da – were periodic by 14 Da. To a certain extent (appendix B) this was verified, and the start of those periodic mass series was summarized in this chart. For each component, a series of maxima correspond to red cells, while minima are blue.

Chromophoric DOM

The four components each occupy a biogeochemically distinct part of the DOM pool. C1 contributes to the majority of the absorbance signal, but comparatively little fluorescence. Thus, this component must represent an important part of the "chromophoric" DOM (CDOM) pool. The masses making up this component are centered between 400 – 900 Da, suggesting that most of the CDOM signal in boreal DOM is not composed of highly humified compounds.

This centering of the CDOM mass range around 600 – 800 Da coincides with mass spectra taken from highly colored black water rivers (i.e., Suwannee River fulvic acid- SRFA), where a shoulder of mass abundance in that spectral region is displayed [*Kujawinski et al.*, 2004]. However, a recent mass spectroscopy study scanning a wider range than ours (0.5 – 4 kDa; *Asmala et al.* [2021]) suggests that CDOM in freshwater systems contributes to m/z far over the 2 kDa measured here. Those compounds, obtained through filtration at 0.7 µm, appear to compensate for their low occurrence by strong absorptivities.

Unsurprisingly, this CDOM dominant component is photoreactive, experiencing significant depletion in all samples exposed to photodegradation. The absorbance spectrum slope of C1 is shallower than the other components, and it has been well documented that photochemical bleaching processes result in a loss of color across the UV-A and a steepening of the spectral slope [*Helms et al.*, 2008; *Reader et al.*, 2015]. In this case, we see a loss of the shallow slope component

with photodegradation, mirroring the results of previous photobleaching studies. Furthermore, this component exhibits a depletion in summer compared to winter, suggesting that the photolysis signature can also be identified in the field. Such behavior of the CDOM pool was previously seen in this Swedish boreal river context by *Reader et al.* [2014], reasserting the CDOM interpretation of C1. Previous studies have also shown this mass range (400 – 900 Da) to be photolabile by subtracting MS fingerprints of natural water from UV-treated samples [*Kujawinski et al.*, 2004; *Stabenau et al.*, 2004].

Interestingly, only 1 ionic series of maxima (periodic at each 14 Da) was identified as a C1 marker (Table 1), which denotes a small compositional diversity producing this chromophoric signal. This corroborates *Kujawinski et al*. [2004] and their ESI-FTICR-MS analysis in which only a highly restricted set of moieties were efficiently targeted by photolysis. Those ions were characterized by specific double bond equivalent values in proportion to low oxygen content, and within an alternate pattern first described by *Stenson et al.* [2003].

Contrasting to C1, C4 had an absorbing peak at 340 nm and was composed of heavier masses (>1 kDa). This recalls the spectral slope ratio property of CDOM as identified by *Helms et al.* [2008], in which high molecular weight compounds were shown to absorb at longer wavelengths. Therefore, the proportion C1/C4 may well mirror the Sr relation to molecular weight, especially as absorption wavelengths of C1 and C4 do correspond to those used to calculate Sr.

Refractory DOM

By contrast, C2 is remarkable for its refractory characteristics as it was found to be stable throughout the year and was unaffected by either photochemical or biological degradation treatments, though it tends to be enriched in the Lyckeby river relative to the other two rivers. C2 contributes to a remarkably constant proportion of the absorbance signal and the majority of the

fluorescence signal at the shorter wavelengths, suggesting that it is part of the ubiquitous fluorescence signal found in most FDOM studies.

As noted earlier, column interactions occur with C2 that tend to make the separation differ from a theoretical SEC behavior where based solely on mass, it should have eluted C3's low masses after C2's. Empirically it did not, which suggests a higher hydrophobicity in C2 than the rest of the DOM pool [*Hawkes et al.*, 2019; *Kellerman et al.*, 2015]. Such views would be consistent with its peak A type of fluorescence ("humic-like"), and Lyckeby's enrichment as the most forested catchment. Acknowledging as well C2's stability to degradation, evidence converges toward a "humic" interpretation of this chemical subset.

Several other characteristics point to C2 as the key component of DOM's ubiquitous refractory background signal. Its mass fingerprint is strikingly similar to those found in marine deep-water DOM for the aged (or refractory) DOM signal [*Flerus et al.*, 2012; *Hansman et al.*, 2015]. Moreover, C2 mass periodicity is very regular with a set of 4 maxima spaced by 2 Da ($+H_2$ transitions and/or (minus CH_2 – plus O)) and alternating progressively along the periods (Appendix B). This shifting periodicity resembles the family of carboxyl-rich alicyclic molecules (CRAM), in which chain elongation is defined by carboxy enrichment (+44 Da).

The deep-sea refractory DOM pool has been widely characterized as having a significant CRAM content, fewer heteroatoms (N, P, and S), the dominance of a diversified set of carbon-based aliphatic, branched ethyl and ethylene groups, and higher ketone and ester moieties in contrasting decreasing alcohol and ether content [*Hertkorn et al.*, 2013]. The cumulative effects of these typically non-bioavailable chemical properties and a high antioxidant capacity (*Catalá et al.* [2020]) explain well the observed resistance to remineralization.

While the common understanding is that refractory DOM is produced in the ocean's interior via the microbial loop [*Jiao et al*., 2010; *Repeta*, 2015; *Shen and Benner*, 2018], our results contrast

by suggesting that there is some fraction of this refractory CRAM-like material produced in surface freshwaters as well. The microbial loop type system likely exists in all aquatic environments, but the freshwater contribution may have been masked until now by methods not capable of separating DOM into subcomponents. If our riverine fingerprint does indeed correspond to the aged marine water one, it would suggest a transport and accumulation of that RDOM from the freshwater system into deep marine water masses, in addition to production within the ocean itself.

Fluorescent DOM

Regarding C3, the FDOM mass fingerprint has some similarities with those presented by *Stubbins et al.* [2014]. Through a different approach to testing the correlation between MS signal and fluorescence peaks intensity, they identify the mass repartition, the aromaticity index, and the N content of those compounds from 0.1 to 1 kDa. For the most part, they confirm the distribution of FDOM from 200 to 500 Da, with an outlier group at 600 Da corresponding to peak A (as defined by *Coble et al.*, [2014]) with low bioavailability and high photosensitivity. Indeed, our C3 fluorescence (λ_{abs}< 280nm, λ_{em}= 450nm) is similar to peak A, and displays similar degradation lability, and bimodality of masses. More recently, those properties elements were again observed and refined by *Zhang et al.* [2020].

Overall it suggests a proportionality of small masses concentrations in the range of monolignols and amino-acid residues with other masses of about 0.6 to 1 kDa. *Stubbins et al*. [2014] explained this as an effect of either: 1) similar environmental sources/sinks, or 2) a cooperative dissolution of light aromatics and aliphatic thanks to smaller polysaccharides and polypeptides. From there and following the postulate of *Simpson et al.* [2002], they propose that heavy “humic-like” FDOM compounds are in fact complexes of fluorescent and non-fluorescent compounds, and so suggest our results as well. Besides all this, a striking distinction happens

between the loadings of maxima (Table 2.1: phase 203 and 205 Da) and minima ion families (phase 209 and 211 Da), which highlights a small set of structures but of various lengths.

Labile DOM

Finally, C4 is sensitive to both photochemistry and biodegradation. With most of its fingerprint >0.8 kDa, C4 is likely composed of a set of biopolymers whose chemical linkages are easily disrupted by photochemistry and microbial nutrient acquisition strategies. The transport of labile chemical moieties in freshwater is a well-known process, characterized by a strong relationship to river discharge [*Messetta et al.*, 2018; *Warrick*, 2015].

Such behavior of DOM concentration derives from fluid mechanics and was adapted by hydrologists for sediment transport, which then proved useful to DOM studies. It assumes a proportionality of particles' content with the change of height in a river, and as soil particle amount increases so does the DOM leaching from these. Their models predict that $[DOM] \approx a \cdot Q^b$, where Q is the normalized discharge, and a and b empirical constants are derived from regression fitting [*Braca*, 2008; *Cohn et al.*, 1992]. The general view is that higher flows induce run-off from soil layers more abundant in organic matter, an effect flattened by the widening of the river canal along with discharge increase [*Butturini et al.*, 2005; *Vidon et al.*, 2008].

Another convincing element that C4 does account for the LDOM fraction of our samples, is its linear positive relationship to BOD in all rivers (overall $p_{value}= 6 \times 10^{-5}$). It implies that C4 represents a major resource fueling biological carbon respiration. This put in perspective the periodicity of C4 mass-loadings, for which there is a clear distinction between minima and maxima. While the ionization method may indeed bias the detections of analytes, such a clear distinction for one of our components denotes a strong metabolic selectivity of certain ion families. Instead of a "harmonic" pattern as for C2 (Appendix B), certain masses were very high and others as low as 0,

which indicates the accommodation of all ions within a selectively chosen homologous family. This microbial specificity was also marked by the non-periodic maximum at 248 Da.

In regards to C4 fluorescence, with emission at 490 nm and with an absorbance peak at 340 nm, it does not match Coble's peak for LDOM [*Coble et al.*, 2014]. However, the experimental set-up with the excitation wavelength (λ_{max}= 254 nm) is likely mismatching the true excitation wavelength, given the absorbance spectrum of the component. Our C4 more closely resembles Component 2 in the study of *Lapierre and del Giorgio* [2014]. They identified an FDOM signature (275 (340), 484) from Canadian boreal aquatic networks that hold the same transition (340 → 484) as our data. However, this "terrestrially derived material" as labeled by them was photolabile but not biodegradable, and their component proportion in the pool (~30% of the fluorescence signal) does not match ours (~3%). Considering the limitations of PARAFAC, it is possible that their FDOM signature represented the cumulative profile of our C1 and C4, which would have then displayed the contrasting properties aforementioned (abundance, transitions, and lability).

A lack of discrimination in FDOM is plausible as even in our results, it was noticed that C4 scores were positively correlated to C1's (Appendix A.D: R^2 = 0.45). This covariance is due to a similar reactivity pattern between the two pools -photodegradable and summer-depleted. However, statistics revealed an insignificant effect of biological activity for C1 (p_{value}= 0.88) unlike C4. Adding to the fact that they do not elute at the same time, represent different MS-spectra, CDOM, and FDOM shapes those pools have differed enough to be dissociated in our PARAFAC model, but not in Lapierre and del Giorgios's study.

Overall, C4 is a subset constituted of labile terrestrial matter following seasonal hydrological conditions. Hence, where C2 captures the RDOM fingerprint, we showed the contrasting behavior of C4 as a key part of the LDOM pool. Yet, we failed to observe any component produced as degradation occurs (Figure 2.3). This might be due to the stochastic nature of chemical degradation,

which tends to create multiple intermediates not necessarily consistent between samples [*Leyva et al.*, 2022], and therefore not modellable under PARAFAC.

2.5.2 DOM seasonal reactivity in the perspective of fingerprints' chemistry

We quantified seasonal variation as the difference between the January and August samples. For C1 the match between the photochemical degradation scores and the winter-summer scores suggests summer depletion due to photosensitivity. This calls to mind the known phenomenon of "winter preservation". Cold temperatures and ice cover, as well as shorter sunlight hours, probably lower degradation rates of C1 within the aquatic systems without impacting release processes, thus inducing higher concentrations in winter [*Ågren et al.*, 2008b; *Haei et al.*, 2012]. Therefore, the chromophoric C1 is likely a pool dynamically linked to temperature conditions across the year.

Despite C3's photolability, its abundance within the pool was not lower during the summer but stayed the same in Emån and Ume and increased in Lyckeby. This suggests a summer process compensating and/or surpassing photolytic losses. Considering that *Reader et al.* [2015] observed a summer production of DOM within a similar sample set, this possibility seems credible. Overall, due to the counteracting action of summer production off-setting photochemical degradation, the seasonal variation of the fluorescing component C3 is low in these Swedish rivers.

Finally, C4 demonstrated semi-labile properties and has a mass fingerprint coherent with the theory of the size-reactivity continuum [*Amon and Benner*, 1996]. According to this perspective, important reprocessing over 0.7 kDa during biodegradation assays is linked to a higher chemical hydrolysability potential [*Stabenau et al.*, 2004; *Xu and Guo*, 2018]. That said the sources of DOM are also critical to fully understanding C4's dynamics. As found by *Messetta et al.* [2018], leachates originate from shallower soil layers during high-flow periods (excluding storm events). This increases the proportion of fresh allochthonous humic compounds, explaining the relation between

the discharge and chemical properties such as the humification index, the biological index, etc. Then, these fresh soil-related compounds were associated with a priming of biological activity when the flow decreased again. Similar findings are observed throughout multiple ecosystems, and in particular in Swedish boreal catchments where the lability of our C4 is characteristic of fresh soil material released during high flow conditions [Ågren et al., 2008a; *Vidon et al.*, 2008].

2.5.3 Spatial-abundance variability related to the freshwater system complexity

Previous findings made in 2015 over the same 3 catchments, concluded that lakes and forest cover determined significant divergence of DOM composition [*Reader et al*., 2015]. Those findings support ours, with spatial differentiation of mass fingerprints in Lyckeby and Ume with respectively a more abundant C2 (CRAM content) and less C3 (humic-like FDOM).

Lyckeby river, the smallest of the 3 catchments with 810 km^2 and a geometric mean flow of 2.6 m^3/s, was significantly more abundant in C2. Per unit of area, Lyckeby drains the least amount of water, and the literature suggests that such low flow conditions favor water originating from hyporheic sediment [*Lynch et al.*, 2019; *Thurman*, 1985]. Thus, a higher proportion of aged and reprocessed organic matter enters the river from the soil, which in the case of the Lyckeby catchment, covered at 72% by the boreal forest, is typically non-labile. Hence, the smaller flow levels of Lyckeby predisposed it to contain a higher level of RDOM. Empirically, C2 scores were extremely stable across the season, as if they originated from a constant baseflow relatively independent of surface variable runoff. C2 could effectively originate from the percolation of deep layers of the soil, levels at which only remains the refractory material from the boreal litter. Those properties enhanced in Lyckeby could explain its enrichment in aromatic content characterized by *Reader et al.* [2014].

On the other hand, Ume, the biggest catchment with 26 800 km^2 and a mean flow of 509 m^3/s, was significantly depleted in C3. In 2013, a large-scale survey of Swedish boreal lakes by *Kothawala et al.* [2014] did observe a depletion trend of humic-like FDOM with water surface area and water residence time in those lakes. Not only does a wider unvegetated area enhance photolytic action, but the calmer flow of lakes also enables more DOM flocculation with sediments (removal by diagenesis). Those propositions fit our observations as 8% of Ume catchment is covered by lakes, which is higher than Emån's (6%; and 2nd lowest C3 scores) and Lyckeby's (4%; highest scores). Thus, even if we cannot strictly rule out other hypotheses as climatological or production variables, water coverage of catchments seems like a critical factor determining the C3 chemical subset in our dataset.

Regarding C4 and unlike behaviors in Emån and Lyckeby, we failed to observe a clear trend between its score in Ume and the riverine discharge. This is likely due to Ume's catchment complexity which covers 6% of Sweden's territory and consists of two large subcatchments, one forested (Vindeln) and the other more open (Ume). Each of those subregions has independent DOM pool behavior, but as the relative contribution of each to the downstream sampling site changes greatly over a year, the integrated signal pattern in Ume is likely blurred by parallel processes unfolding asynchronously. Ume is a highly regulated river, transformed by infrastructure like dams which affect the sediment transport and reprocessing within the system [*Berglund*, 2014]. Hence, this blurring of a pattern between discharge and fresh terrestrial input in Ume could be related to the complexity of the catchment compared to the two southern rivers. Another element could be that with the spring freshet, a high proportion of the stocks of LDOM from the soil surface is washed away, leaving little amounts to percolate for the rest of the summer until biomass decay in the fall. In such a situation where C4 stocks turn into a limiting factor, the discharge would be relegated to a smaller impact than the seasonal availability.

2.5.4 PARAFAC as an approach for DOM fingerprinting

This approach successfully modeled a dataset containing an elution-time dimension. Considering, that it was formed by roughly 34 $\times 10^6$ data points, keeping the tridimensional structure enables less computational effort than its slicing (dimensional reduction of the matrix by its reorganization: e.g., $5\times4\times3 \rightarrow 5\times12$), which would have also ruined the consistency between variables profiles [*Bro et al.*, 2010]. On the other hand, we had to constrain the model with nonnegativity assumptions to avoid degeneration. This hindered the identification of production behavior within components. If a compound degraded into another ($m_1 \rightarrow m_2$), a non-constrained model could have represented a negative score for m_1 and a positive for m_2 in the same component. Although such true connections would be quite more complex, here, the non-negativity constraint prevents their identification anyway. Thus, it failed to observe any production pattern associated with DOM reprocessing within components, though this does not necessarily prohibit production signals caused by shifting compounds between components (i.e., summer production in C3).

Surprisingly, no more than 4 components were yielded along 2185 variables and 190 retention time points per sample. This corroborates *Wünsch and Hawkes* [2020] observations where a low number of components (N=5) contrasted with the very high-resolution MS data. This seems to be due to PARAFAC's underlying assumptions which are efficient in co-clustering quantity variations [*Murphy et al.*, 2013]. Hence, just as worldwide FDOM patterns seem to be explained by only 7 components under PARAFAC [*Jørgensen* et al., 2011], mass spectrometry data may be shaped by a low number of parameters through this approach.

Depending on the study objectives, such a parsimonious model can be the desired strategy to highlight compounds' behavioral similarities, or not. In certain contexts of analytical chemistry like *Farag et al.* [2018], the aim is to yield as many unique biomarkers as possible, so a PARAFAC fingerprinting approach does not seem adequate. However, it is a good strategy for environmental

studies interested in DOM fingerprints, whether mass or optical spectra. Here, fingerprints include the entire set of compounds sensitive to a particular condition (source, reactivity, etc). The clustering strategy behind fingerprinting approaches has already been defended as identifying more meaningful trends than focusing on individual metabolite patterns [*Alonso-Sáez et al.*, 2012; *Amaral et al.*, 2016].

At first glance at DOM complexity, it seems plausible that each m/z has its independent behavior in the environment. But the accumulating set of evidence of a DOM pool described by only a small number of components [*Dąbrowski*, 2020; *Wünsch et al*., 2017], indicates that compounds' proportion is closely internally linked. To a certain extent, this might be associated with chemical equilibrium constants and pseudo steady-state production and degradation rates. However, how this inter-dependence evolves through spatio-temporal variations as well as microscale patchiness is driven by a yet undescribed set of mechanisms [*Stocker*, 2012].

Conclusion

The cumulative set of evidence indicates that we identified fingerprints for key DOM sub-pools based on their mass spectra, optical properties, and compositional lability. Terminology as “chromophoric” DOM, “fluorescent DOM”, “refractory DOM”, and “labile DOM” are often used in the literature without a clear understanding of the extent of their superposition. Here we provide the cumulative averaged abundance over the mass dimension unraveling how seasonal changes remodel the pool MS-spectrum shape. The information we gathered from an elution smear was coherent with most of the spatial and temporal variations described in the literature.

It is unknown how well the mass-periodic maxima and minima phases we identified apply in other contexts. Yet if replication of those within similar ecosystems (boreal freshwater) and over

multiple annual cycles was confirmed, it could represent a handy tool to bridge the correspondence of effects between routine DOM optical properties and DOM mass distribution.

Finally, chemometric fingerprinting of DOM here again demonstrated its efficiency to detect deep-set trends of DOM remodeling. This approach produces parsimonious data explanations, which suggest mechanisms interrelating the proportions of different compounds. The balance between those abundances represents an underexplored theoretical field of biogeochemistry, which could help to bridge the gaps between cell metabolome and DOM behavior.

2.6 Reference

Abdulla, H., E. C. Minor, R. F. Dias, and P. G. Hatcher (2010), Changes in the compound classes of dissolved organic matter along an estuarine transect: A study using FTIR and 13C NMR, *Geochimica et Cosmochimica Acta*, *74*(13), 3815-3838, doi:https://doi.org/10.1016/j.gca.2010.04.006.

Ågren, A., M. Berggren, H. Laudon, and M. Jansson (2008a), Terrestrial export of highly bioavailable carbon from small boreal catchments in spring floods, *Freshwater Biology*, *53*(5), 964-972, doi:https://doi.org/10.1111/j.1365-2427.2008.01955.x.

Ågren, A., I. Buffam, M. Berggren, K. Bishop, M. Jansson, and H. Laudon (2008b), Dissolved organic carbon characteristics in boreal streams in a forest-wetland gradient during the transition between winter and summer, *Journal of Geophysical Research: Biogeosciences*, *113*(G3), doi:https://doi.org/10.1029/2007JG000674.

Alonso-Sáez, L., O. Sánchez, and J. M. Gasol (2012), Bacterial uptake of low molecular weight organics in the subtropical Atlantic: Are major phylogenetic groups functionally different?, *Limnology and Oceanography*, *57*(3), 798-808, doi:https://doi.org/10.4319/lo.2012.57.3.0798.

Amaral, V., D. Graeber, D. Calliari, and C. Alonso (2016), Strong linkages between DOM optical properties and main clades of aquatic bacteria, *Limnology and Oceanography*, *61*(3), 906-918, doi:10.1002/lno.10258.

Amon, R., and R. Benner (1996), Bacterial utilization of different size classes of dissolved organic matter, *Limnology and Oceanography*, *41*(1), 41-51, doi:https://doi.org/10.4319/lo.1996.41.1.0041.

Asmala, P. Massicotte, and J. Carstensen (2021), Identification of dissolved organic matter size components in freshwater and marine environments, *Limnology and Oceanography*, *66*(4), 1381-1393, doi:https://doi.org/10.1002/lno.11692.

Berglund, L. (2014), Riparian vegetation distribution along the Ume river : Predicted responses of riparian plants to environmental flow modifications in run-of-river impoundments (Dissertation) edited, Umeå University, Department of Ecology and Environmental Sciences. http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-89687.

Braca, G. (2008), *Stage-discharge relationships in open channels: pratices and problems - FORALPS Technical Report, 11*, Università degli Studi di Trento, ISBN 978–88–8443–230–8.

Bro, R. (1997), PARAFAC. Tutorial and applications, *Chemometrics and Intelligent Laboratory Systems*, *38*(2), 149-171, doi:https://doi.org/10.1016/S0169-7439(97)00032-4.

Bro, R., N. Viereck, M. Toft, H. Toft, P. I. Hansen, and S. B. Engelsen (2010), Mathematical chromatography solves the cocktail party effect in mixtures using 2D spectra and PARAFAC, *Trends in Analytical Chemistry*, *29*(4), 281-284, doi:https://doi.org/10.1016/j.trac.2010.01.008.

Butturini, A., S. Bernal, and F. Sabater (2005), Modeling storm events to investigate the Influence of the stream-catchment interface zone on stream biogeochemistry, *Water Resources Research*, *41*(8), 12 p., doi:10.1029/2004WR003842.

Bylund, D., R. Danielsson, G. Malmquist, and K. E. Markides (2002), Chromatographic alignment by warping and dynamic programming as a pre-processing tool for PARAFAC modelling of liquid chromatography-mass spectrometry data, *J Chromatogr A*, *961*(2), 237-244, doi:10.1016/s0021-9673(02)00588-5.

Carlson, C. A., and D. A. Hansell (2015), Chapter 3 - DOM sources, sinks, reactivity, and budgets, in *Biogeochemistry of marine dissolved organic matter (second edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 65-126, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00003-0.

Catalá, T. S., P. E. Rossel, F. Álvarez-Gómez, J. Tebben, F. L. Figueroa, and T. Dittmar (2020), Antioxidant activity and phenolic content of marine dissolved organic matter and their relation to molecular composition, *Frontiers in Marine Science*, *7*(984), 17 p., doi:10.3389/fmars.2020.603447.

Coble, P. G., J. Lead, A. Baker, D. M. Reynolds, and R. G. M. Spencer (2014), *Aquatic organic matter fluorescence*, Cambridge University Press, Cambridge, doi:10.1017/CBO9781139045452.

Cohn, T. A., D. L. Caulder, E. J. Gilroy, L. D. Zynjuk, and R. M. Summers (1992), The validity of a simple statistical model for estimating fluvial constituent loads: An Empirical study involving nutrient loads entering Chesapeake Bay, *Water Resources Research*, *28*(9), 2353-2363, doi:https://doi.org/10.1029/92WR01008.

Cory, R. M., and G. W. Kling (2018), Interactions between sunlight and microorganisms influence dissolved organic matter degradation along the aquatic continuum, *Limnology and Oceanography Letters*, *3*(3), 102-116, doi:https://doi.org/10.1002/lol2.10060.

Dąbrowski, Ł. (2020), Evaluation of a simplified method for GC/MS qualitative analysis of polycyclic aromatic hydrocarbons, polychlorinated biphenyls, and organic pesticides using PARADISe computer program, *Molecules*, *25*(16), 3727, doi:10.3390/molecules25163727.

del Giorgio, P. A., and M. L. Pace (2008), Relative independence of organic carbon transport and processing in a large temperate river: The Hudson River as both pipe and reactor, *Limnology and Oceanography*, *53*(1), 185-197, doi:https://doi.org/10.4319/lo.2008.53.1.0185.

Denis, D. J. (2020), Chapter 9- Multivariate analysis of variance (MANOVA) and discriminant analysis, in *Univariate, bivariate, and multivariate statistics using R*, edited, pp. 251-279, doi:https://doi.org/10.1002/9781119549963.

Derrien, M., S. R. Brogi, and R. Goncalves-Araujo (2019), Characterization of aquatic organic matter: Assessment, perspectives and research priorities, *Water Res*, *163, 114908*, 17 p., doi:10.1016/j.watres.2019.114908.

Farag, M. A., A. Meyer, S. E. Ali, M. A. Salem, P. Giavalisco, H. Westphal, and L. A. Wessjohann (2018), Comparative Metabolomics Approach Detects Stress-Specific Responses during Coral Bleaching in Soft Corals, *J Proteome Res*, *17*(6), 2060-2071, doi:10.1021/acs.jproteome.7b00929.

Flerus, R., O. Lechtenfeld, B. Koch, S. McCallister, P. Schmitt-Kopplin, R. Benner, K. Kaiser, and G. Kattner (2012), A molecular perspective on the ageing of marine dissolved organic matter, *Biogeosciences*, *9*(6), 1935–1955, doi:10.5194/bg-9-1935-2012.

Gibson, J., and C. S. Harwood (2002), Metabolic Diversity in Aromatic Compound Utilization by Anaerobic Microbes, *Annual Review of Microbiology*, *56*(1), 345-369, doi:10.1146/annurev.micro.56.012302.160749.

Goldman, J. C., D. A. Caron, and M. R. Dennett (1987), Regulation of gross growth efficiency and ammonium regeneration in bacteria by substrate C: N ratio, *Limnology and Oceanography*, *32*(6), 1239-1252, doi:https://doi.org/10.4319/lo.1987.32.6.1239.

Haei, M., M. G. Öquist, U. Ilstedt, and H. Laudon (2012), The influence of soil frost on the quality of dissolved organic carbon in a boreal forest soil: combining field and laboratory experiments, *Biogeochemistry*, *107*(1), 95-106, doi:10.1007/s10533-010-9534-2.

Hansell, D. A., and C. A. Carlson (2014), *Biogeochemistry of marine dissolved organic matter*, 698 pp., Elsevier Science & Technology, San Diego, US, ISBN 9780124071537.

Hansman, R. L., T. Dittmar, and G. J. Herndl (2015), Conservation of dissolved organic matter molecular composition during mixing of the deep water masses of the northeast Atlantic Ocean, *Marine Chemistry*, *177*, 288-297, doi:https://doi.org/10.1016/j.marchem.2015.06.001.

Hawkes, J. A., et al. (2020), An international laboratory comparison of dissolved organic matter composition by high resolution mass spectrometry: Are we getting the same answer?, *Limnology and Oceanography: Methods*, *18*(6), 235-258, doi:https://doi.org/10.1002/lom3.10364.

Hawkes, J. A., P. J. R. Sjöberg, J. Bergquist, and L. J. Tranvik (2019), Complexity of dissolved organic matter in the molecular size dimension: insights from coupled size exclusion chromatography electrospray ionisation mass spectrometry, *Faraday Discussions*, *218*(0), 52-71, doi:10.1039/C8FD00222C.

Helms, J. R., A. Stubbins, J. D. Ritchie, E. C. Minor, D. J. Kieber, and K. Mopper (2008), Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter, *Limnology and Oceanography*, *53*(3), 955-969, doi:10.4319/lo.2008.53.3.0955.

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*(3), 1583-1624, doi:10.5194/bg-10-1583-2013.

Holm, S. (1979), A simple sequentially rejective multiple test procedure, *Scandinavian Journal of Statistics*, *6*(2), 65-70.

Jiao, N., et al. (2010), Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean, *Nature Reviews Microbiology*, *8*(8), 593-599, doi:10.1038/nrmicro2386.

Jørgensen, L., C. A. Stedmon, T. Kragh, S. Markager, M. Middelboe, and M. Søndergaard (2011), Global trends in the fluorescence characteristics and distribution of marine dissolved organic matter, *Marine Chemistry*, *126*(1), 139-148, doi:https://doi.org/10.1016/j.marchem.2011.05.002.

Kellerman, A. M., D. N. Kothawala, T. Dittmar, and L. J. Tranvik (2015), Do we have a weight problem? The discrepancy between actual and apparent molecular weight of dissolved organic matter in lakes -manuscript, in *Thesis...Molecular-level dissolved organic matter dynamics in lakes: Constraints on reactivity and persistence*, edited, University Uppsala, Department of Ecology and Genetics.

Kothawala, D. N., C. A. Stedmon, R. A. Müller, G. A. Weyhenmeyer, S. J. Köhler, and L. J. Tranvik (2014), Controls of dissolved organic matter quality: evidence from a large-scale boreal lake survey, *Global Change Biology*, *20*(4), 1101-1114, doi:https://doi.org/10.1111/gcb.12488.

Kujawinski, E. B., R. Del Vecchio, N. V. Blough, G. C. Klein, and A. G. Marshall (2004), Probing molecular-level transformations of dissolved organic matter: insights on photochemical degradation and protozoan modification of DOM from electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, *Marine Chemistry*, *92*(1), 23-37, doi:https://doi.org/10.1016/j.marchem.2004.06.038.

Lapierre, J. F., and P. A. del Giorgio (2014), Partial coupling and differential regulation of biologically and photochemically labile dissolved organic carbon across boreal aquatic networks, *Biogeosciences*, *11*(20), 5969-5985, doi:10.5194/bg-11-5969-2014.

Lechtenfeld, O. J., N. Hertkorn, Y. Shen, M. Witt, and R. Benner (2015), Marine sequestration of carbon in bacterial metabolites, *Nat Commun*, *6*(1), 6711, doi:10.1038/ncomms7711.

Leyva, D., M. U. Tariq, R. Jaffé, F. Saeed, and F. F. Lima (2022), Unsupervised structural classification of dissolved organic matter based on fragmentation pathways, *Environmental Science & Technology*, *56*(2), 1458-1468, doi:10.1021/acs.est.1c04726.

Lynch, L. M., N. A. Sutfin, T. S. Fegel, C. M. Boot, T. P. Covino, and M. D. Wallenstein (2019), River channel connectivity shifts metabolite composition and dissolved organic matter chemistry, *Nature Communications*, *10*(1), 459, doi:10.1038/s41467-019-08406-8.

Messetta, M. L., C. Hegoburu, J. P. Casas-Ruiz, A. Butturini, and C. Feijoó (2018), Characterization and qualitative changes in DOM chemical characteristics related to hydrologic conditions in a Pampean stream, *Hydrobiologia*, *808*(1), 201-217, doi:10.1007/s10750-017-3422-x.

Murphy, K. R., C. A. Stedmon, D. Graeber, and R. Bro (2013), Fluorescence spectroscopy and multi-way techniques: PARAFAC, *Analytical Methods*, *5*(23), 6557-6566, doi:10.1039/C3AY41160E.

Murphy, K. R., S. A. Timko, M. Gonsior, L. C. Powers, U. J. Wünsch, and C. A. Stedmon (2018), Photochemistry illuminates ubiquitous organic matter fluorescence spectra, *Environmental Science & Technology*, *52*(19), 11243-11250, doi:10.1021/acs.est.8b02648.

Reader, H. E., and W. L. Miller (2012), Variability of carbon monoxide and carbon dioxide apparent quantum yield spectra in three coastal estuaries of the South Atlantic Bight, *Biogeosciences*, *9*(11), 4279-4294, doi:10.5194/bg-9-4279-2012.

Reader, H. E., C. A. Stedmon, and E. S. Kritzberg (2014), Seasonal contribution of terrestrial organic matter and biological oxygen demand to the Baltic Sea from three contrasting river catchments, *Biogeosciences*, *11*(12), 3409-3419, doi:10.5194/bg-11-3409-2014.

Reader, H. E., C. A. Stedmon, N. J. Nielsen, and E. S. Kritzberg (2015), Mass and UV-visible spectral fingerprints of dissolved organic matter: sources and reactivity, *Frontiers in Marine Science*, *2*(88), 10 p., doi:10.3389/fmars.2015.00088.

Repeta, D. J. (2015), Chapter 2 - Chemical characterization and cycling of dissolved organic matter, in *Biogeochemistry of marine dissolved organic matter (second edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 21-63, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00002-9.

Shen, Y., and R. Benner (2018), Mixing it up in the ocean carbon cycle and the removal of refractory dissolved organic carbon, *Scientific Reports*, *8*(1), 2542, doi:10.1038/s41598-018-20857-5.

Simpson, A. J., W. L. Kingery, M. H. Hayes, M. Spraul, E. Humpfer, P. Dvortsak, R. Kerssebaum, M. Godejohann, and M. Hofmann (2002), Molecular structures and associations of humic substances in the terrestrial environment, *Naturwissenschaften*, *89*(2), 84-88, doi:10.1007/s00114-001-0293-8.

Stabenau, E., R. Zepp, E. Bartels, and R. Zika (2004), Role of the seagrass Thalassia testudinum as a source of chromophoric dissolved organic matter in coastal south Florida, *Marine Ecology-progress Series*, *282*, pp. 59-72, doi:10.3354/meps282059.

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, *Analytical Chemistry*, *75*(6), 1275-1284, doi:10.1021/ac026106p.

Stocker, R. (2012), Marine microbes see a sea of gradients, *Science*, *338*(6107), 628-633, doi:10.1126/science.1208929.

Stubbins, A., J. F. Lapierre, M. Berggren, Y. T. Prairie, T. Dittmar, and P. A. del Giorgio (2014), What's in an EEM? Molecular signatures associated with dissolved organic fluorescence in boreal Canada, *Environ Sci Technol*, *48*(18), 10598-10606, doi:10.1021/es502086e.

Thurman, E. M. (1985), Chapter 2- Transport, origin and source of dissolved organic carbon, in *Organic geochemistry of natural waters*, edited, pp. 67-85, Vol. 2. 1st edition. Springer Netherlands, Dordrecht, doi:10.1007/978-94-009-5095-5_3.

Vidon, P., L. E. Wagner, and E. Soyeux (2008), Changes in the character of DOC in streams during storms in two midwestern watersheds with contrasting land uses, *Biogeochemistry*, *88*(3), 257-270, URL: https://www.jstor.org/stable/40343578.

Warrick, J. A. (2015), Trend analyses with river sediment rating curves, *Hydrological Processes*, *29*(6), 936-949, doi:https://doi.org/10.1002/hyp.10198.

Wünsch, U. J., and J. A. Hawkes (2020), Mathematical chromatography deciphers the molecular fingerprints of dissolved organic matter, *Analyst*, *145*(5), 1789-1800, doi:10.1039/C9AN02176K.

Wünsch, U. J., K. R. Murphy, and C. A. Stedmon (2017), The one-sample PARAFAC approach reveals molecular size distributions of fluorescent components in dissolved organic matter, *Environmental Science & Technology*, *51*(20), 11900-11908, doi:10.1021/acs.est.7b03260.

Xu, H., and L. Guo (2018), Intriguing changes in molecular size and composition of dissolved organic matter induced by microbial degradation and self-assembly, *Water Research*, *135*, 187-194, doi:https://doi.org/10.1016/j.watres.2018.02.016.

Zhang, X., J. Kang, W. Chu, S. Zhao, J. Shen, and Z. Chen (2020), Spectral and mass spectrometric characteristics of different molecular weight fractions of dissolved organic matter, *Separation and Purification Technology*, *253*(117390), 10 p., doi:https://doi.org/10.1016/j.seppur.2020.117390.

Zhuang, W.-E., and L. Yang (2018), Impacts of global changes on the biogeochemistry and environmental effects of dissolved organic matter at the land-ocean interface: a review, *Environmental Science and Pollution Research*, *25*(5), 4165-4173, doi:10.1007/s11356-017-1027-6.

Chapter 3: Diverging chemical and optical properties of DOM during the growth of coastal phytoplankton assemblages

Manuscript

Authors

Antoine V. Viallard [1]; Rachel E. Sipler[2, 3]; and Heather E. Reader [4]

Affiliations

[1] Environmental Science Program, Memorial University of Newfoundland, St John's (NL), Canada

[2] Ocean Science Center, Memorial University of Newfoundland, St. John's (NL), Canada

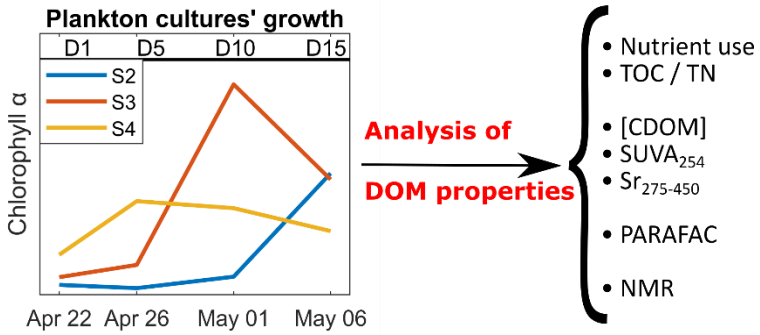
[3] Bigelow Laboratory for Ocean Sciences, East Boothbay (ME), USA

[4] Department of Chemistry, Memorial University of Newfoundland, St John's (NL), Canada

3.1 Co-authorship statement and acknowledgment:

Acknowledgment: The experimental design was an original idea of HR and RS, developed and performed by AVV. Nutrient measurements were taken by Maria Ignacia Diaz at the Ocean Science Center (NL, Canada). NMR measurements were performed by Céline Schneider through C-CART facilities at the Memorial University of Newfoundland. The manuscript was written by AVV and benefited from HR's input as well as those of the thesis committee (Rachel Sipler and Uta Passow). This current version is provided under the approval of all intellectual property owners.

Financial support: The experiment was supported by the Natural Science and Engineering Research Council of Canada with the CGS M-2021 program allocated to AVV, and the NSERC Discovery Grant and Canada Research Chair Programs to HR. The cost of nutrient measurements was covered by the Ocean Frontier Institute (project name: *The Northwest Atlantic as a Climate Ocean Projecting Future Changes in Productivity and the Biological Carbon Pump*).



Abstract and graphical abstract

The chemical characteristics of the dissolved organic matter (DOM) in coastal waters are in good part defined by the exudates of the phytoplankton developing in spring. Yet, several classes of plankton species coexist in the community (e.g., diatoms, bacteria, etc.), and we hypothesized that they do not generate the same imprint in the DOM pool. Such chemical properties are also changing across bloom stages. Hence, we investigated the succession of those DOM characteristics to develop our understanding of carbon cycling in boreal regions like the island of Newfoundland.

In this experiment, a series of fractionated microbial incubations were conducted from phytoplankton collected in an estuary of Placentia Bay. As growth progressed, a wide set of DOM chemical and optical proxies were monitored, unraveling the relationship between growth stages and exudates composition. Whereas ~85% of DOC exudation occurred to be in the carbohydrate-related band of ^{1}H-NMR, a richer set of evidence from the fluorescent DOM (FDOM) composition suggests that diatom growth remodels each subcomponent in a unique pattern.

Between treatments, we observed contrasting rates of inorganic nutrient consumption and diverging chromophoric properties (CDOM). DOM properties changed considerably during the diatoms' decline phase, a period of nutrient limitation and cell death, such as the experiment highlighted the effect of autotrophic activity and secondary metabolism on DOM imprint.

3.2 Introduction

Dissolved organic matter (DOM) refers to the natural pool of organic compounds found in the water column. Worldwide, it represents 662 Pg of carbon whose composition and properties are still widely unknown [*Catalá et al*., 2021; *Hedges*, 2002]. In coastal areas, most of the DOM comes from the riverine inputs and the exudates of phytoplankton species, the latter peaking during plankton bloom in spring[*Behrenfeld and Boss*, 2018]. The microbial community is diverse, and the DOM characteristics of each class are not similar (e.g., phototrophs vs heterotrophs, bacteria vs diatoms) [*Brown*, 1991; *Kinsey et al.*, 2018; *Larsson and Hagström*, 1982].

Exudates from plankton strains of the North Atlantic have been analyzed since the 1960s testing the metabolic behavior of isolated plankton species under several conditions [*Harrison et al.*, 1977; *Lewin and Lewin*, 1960; *Peltomaa et al*., 2019]. Yet, DOM composition is not simply the sum of exudates from isolated strains as microbes interact metabolically with each other [*Shibl et al.,* 2020; *Spice and Ackers*, 1992]. For that reason, the assessment of DOM from non-isolated cultures remains crucial to accurately represent the plankton-derived DOM characteristics.

In addition to the variable of species assemblage, the biological imprint in DOM changes across the plankton growth stage [*Lynch et al.*, 2019]. Cells not only release substrates through active processes and incidental leaching, but they also absorb and metabolize the DOM, impacting its composition [*Jiao et al.*, 2010; *Shen and Benner*, 2018]. Such a continuous cycle of release and reprocessing drives therefore the DOM composition along with the changes in cellular activity occurring during plankton bloom [*Myklestad*, 1995; *Richardson and Cullen*, 1995].

Following all this, our research investigated the difference in the DOM properties (also termed signature) from a bidimensional experimental design of different growth stages and plankton groups. For the latter, plankton species size is a critical parameter to study in coastal Newfoundland as microbiological records from the *Continuous Plankton Reporter program* [1958-

2020] reveal a trend towards a bigger proportion of cyanobacteria during spring blooms (relative to diatoms abundance). Such effect is likely caused by climate change where warmer water in coastal Newfoundland favors multiple ecological niches of bacteria (both phototroph and heterotroph) [*Head and Sameoto*, 2007; *Rivkin et al.*, 1996]. Diatoms and dinoflagellates thus decline proportionally in the community, such as studies broadly concluded that smaller organisms are favored [*Buck et al.*, 1996; *Daniels et al.*, 2015; *Hinder et al.*, 2012], with a shift in the balance between primary production and heterotrophy [*Coutinho et al.*, 2015; *Figueroa*, 2016].

Normally, diatoms release significant amounts of labile substrates in late growth stages which stimulates the growth of heterotrophs. The heterotrophs reprocess the DOM, and a more complex molecular pool tends to emerge from heterotrophs' secondary metabolism [*Carlson and Hansell*, 2015]. However, little is known about the coastal Newfoundland plankton metabolome and its influence on DOM dynamics. Moreover, modified growth conditions of the microbial communities (space, time, growth intensity, and taxonomic composition) under the effects of climate change may affect the properties of the DOM.

We hypothesize that different classes of plankton size from coastal Newfoundland would produce DOM with different chemical characteristics; and that such changes would be dynamic across the growth of each treatment. To test this, we monitored DOM exudates from three fractions of the plankton community at four different times across their growth phase. While nutrients were consumed and DOC produced, DOM characteristics were assessed via several optical proxies such as CDOM absorption ($a_{350\ nm}$), the specific UV-absorbance at 254 nm ($SUVA_{254}$), and the CDOM spectral slope ratio (Sr). These are indicators of chromophoric DOM (CDOM) concentration, aromatic content [*Weishaar et al.*, 2003], and mass distribution [*Helms et al.*, 2008], respectively. Quantitative ^{1}H-NMR was also assessed as a tool to determine the enrichment or consumption of

different kinds of functional groups in the DOM pool across carbon accumulation [*Hertkorn et al.*, 2013].

Additionally, the fluorescent properties of the DOM (FDOM) were studied quantitatively using parallel factor analysis (PARAFAC) [*Murphy et al.*, 2013]. Although FDOM represents only a small subset of the chemical pool, it has been proven a useful tool for monitoring sources and processing dynamics in the pool, with success in distinguishing autochthonous production from the allochthonous background [*Amaral et al*., 2016; *Coble et al.*, 2014; *Gao and Guéguen*, 2017]. Altogether, this diverse set of DOM proxies aimed to characterize the contribution of various size-fraction to the chemical coastal fingerprints across growth phases. We found that carbohydrate-related compounds constituted most of the exudates while five FDOM components were shifting.

3.3 Material and methods

3.3.1 Experimental set-up and plankton treatments

Sampling was performed at Garden Cove in Placentia Bay, Newfoundland (47.851079˚N; -54.158716˚E; 5m in depth) on April 21st, 2021. There was a phytoplankton bloom with ~9 mg/m^3 of chlorophyll α (Aquapen, Quibit), which was dominated by diatoms (Appendix C). *In situ* water conditions were measured using an Orion Star™ A329 water (Thermo Scientific). Salinity was 23.4 psu, the temperature was 3˚C, and the water was oxygen saturated.

A total of 22 L of this coastal water was collected into a seawater pre-conditioned carboy using a 5 L Niskin bottle (General Oceanics). A 400 mL plankton sample was also gathered by slowly pulling a hand-held plankton net (20 µm, Aquatic BioTechnology S.L.) vertically in the water column from 7 m depth to the surface (three times). Both the water and the concentrated plankton sample were kept cold (4°C) in the dark until processing in the laboratory (~4 hours).

The plankton size-fractionation consisted of a series of filtrations to generate 3 experimental treatments from the untreated coastal water. All treatments were produced in independent triplicates of 1.7 L each, totaling 9 bottles (polycarbonate; acid-washed before use) plus positive and negative controls (N=2x3 bottles). All the equipment was acid-washed and ashed for 4 hours at 450 ˚C when possible. Filtration through 100 µm and 20 µm was achieved by gravity with a Nitex mesh placed in a funnel, while the one at 0.7 µm was carried under vacuum (-300 mbar) with GF/F filters (Whatman, 0.7 µm).

Treatments: S2 consisted of the filtrate of a two-step process at 20 µm and then 0.7 µm, and was meant to contain bacteria only. Treatment S3 was the filtrate of a single filtration at 20 µm and forming therefore the plankton subset under 20 µm. S4 was processed as S2 but was inoculated with 60 mL of the plankton sample collected with the net and pre-filtered at 100 µm. S4 aimed to represent the plankton size fraction between 20 µm and 100 µm. Positive and negative controls

were respectively filtered at 100 µm and 0.2 µm, to contain either the entire plankton community or almost none of it. Later, the samples were checked using light microscopy to verify if the system succeeded in size-fractionating the initial plankton community.

Preparations of the treatments were done on the same day as sample collection, and we let the community rest overnight in the dark at 4˚C. Then, we spiked each treatment with inorganic nutrients up to the normal spring concentrations in the region (final concentrations of 2.6 µM of KH_2PO_4 – 26 µM of Na_2SiO_3 - 14 µM of NH_4Cl -13 µM of KNO_3). Incubation was conducted at 4˚C, with a 14:10 light cycle of 1.84 mW/cm^2 produced by six growth lamps (T5HO: 17 W-45 cm). PAR-centered low-UV growth lamps and polycarbonate bottles limited the photodegradation of DOM. Twice a day, bottles were agitated, and their position was randomized in the incubator.

Growth was followed for 15 days by measuring the *in vivo* fluorescence every two days using the Aquapen (Qubit). To do so, bottles were agitated and 3 mL of the culture was poured into a plastic cuvette without dark adaptation. Then, fluorescence intensity after 50 µs (F_0) was measured using a photon pulse of 3,000 $\mu mol.m^{-2}.s^{-1}$ at λ_{em}= 455 nm. Values of the 10 first readings were averaged to obtain the final F_0 of the day for that sample.

3.3.2 Harvesting DOM samples

Four DOM profiles were sampled during the 15 days incubations. We were aiming to capture DOM composition at the beginning (reference) and during the early growth stage, the end of growth, and the end of the plateau. Using the aquapen data as a guide, we chose to harvest on days 1, 5, 10, and 15, which will be termed D1, D5, D10, and D15 for more simplicity.

The harvesting started by collecting 400 mL of each treatment and keeping those subsamples cold and dark until filtration (maximum of three hours). 4 mL of the sample was preserved with paraformaldehyde 0.25% (w/v) and stored at -20˚C for later microscopic analysis [*Marie et al.*,

2001]. Then, the rest of the sample (~396 mL) was filtered through 0.3 µm (GF-75 Advantec, ashed) under a vacuum (150-300 mbar) to isolate its DOM. The filters were stored frozen and later used for chlorophyll α quantification (conserved at -20˚C, dark) whereas the filtrate was processed further for nutrient analysis and DOM characterization. Filter towers were changed between treatments (S2, S3, S4), and rinsed many times with Milli-Q water between replicates.

From this sample filtration step, the 390 mL filtrate was split for analysis. 40 mL were stored cold and dark for optical properties (i.e., CDOM and FDOM), and 40 mL were preserved at -20˚C for nutrient analysis. The rest was acidified down to pH=2 using 1.6 mL of HCl 2M, and 40 mL of that sample was kept for DOC analysis (frozen at -20˚C) leaving 270 mL for solid-phase extraction (SPE; stored at 4˚C).

Optical properties were measured as quickly as possible following the harvest with a maximum storage of 11 days for absorbance and 8 days for fluorescence. Similarly, SPE was conducted within 9 days following each harvest. Chlorophyll filter content was measured once at the end of the experiment.

3.3.3 Chlorophyll α and microscopical analysis

Preserved samples were checked microscopically at x100 magnification. 1 mL of the preserved microbial suspension was poured into a plastic Sedgewick Rafter, and let settle for 30 minutes. Diatom taxa (> 4 µm) were identified using the online tool Phycokey [*Baker*, 2012]. Photos of the plankton diversity were kept for each sample.

Chlorophyll α content on the filters was measured using the trichromatic method as described by *Jeffrey and Humphrey* [1975] and reviewed by *Aminot and Rey* [2002]. In brief, chlorophyll was extracted in 90% acetone, and its absorbance spectrum (580-800 nm) was acquired with an

Ocean Insights HDX spectrophotometer system. From the relative absorbances at 750, 664, 647, and 630 nm, the chlorophyll α was quantified and converted into a concentration.

3.3.4 Elemental analysis

Total dissolved organic carbon (DOC) and total dissolved nitrogen (TDN) were measured using a Shimadzu TOC-L CPH equipped with a TNM-L nitrogen detector. Samples were analyzed as NPOC (non-purgeable organic carbon), and the instrument was calibrated with acetanilide standards (Acros Organics, recrystallized). Consensus reference material (sea surface Hansell from RSMAS, University of Miami) and an in-house seawater reference were used to assess the stability of the instrument over the run.

Dissolved inorganic nutrients such as silicates, orthophosphate, ammonium, and nitrate plus nitrite (NO_x) were measured using a Lachat 8500. Each sample was processed in triplicate using the USEPA-approved methods for *LACHAT instruments* [2021]. That is #31-114-27-1-A for SiO_x (detection limit (DL) = 8 x10^{-8} M); #31-115-01-1-J for PO_x (DL = 2 x10^{-8} M); #10-107-06-2-O for NH_4^+ (DL = 2 x10^{-7} M); and #31-107-04-1-D for NO_x (DL = 2 x10^{-8} M). For each, a standard curve with at least 5 points was made before analysis. Reactants and standards were obtained from Fisher and Sigma, and the software Omnion 4.0 was used to operate the Lachat. Nutrient data were imported into Matlab, and nutrient consumption was estimated by subtracting their concentrations at the end of the growth from the starting one.

3.3.5 Optical properties

An absorbance spectrum (200-800 nm) was taken for each sample using a Cary 300 UV-VIS instrument (Agilent) using quartz 10 cm cuvettes, with ultrapure water as a blank. The data were imported into Matlab and a baseline correction was applied by subtracting the average signal

between 700 nm and 800 nm, where no CDOM absorption is expected. Changes in $a_{350\ nm}$ were studied as well as for $SUVA_{254}$ [*Weishaar et al.*, 2003]; and the spectral slope ratio (Sr) [*Helms et al.*, 2008]. The latter was calculated using an exponential fitting script (R^2>0.97), with Sr being the ratio of slope from 275–295 nm over the slope from 300-450 nm.

Excitation-Emission fluorescence matrices (EEMs) were taken – λ_{ex}= 220-450 nm and λ_{em}= 280-600 nm – using a Cary Eclipse Fluorescence Instrument (Agilent), with an increment of 5 nm in excitation and 2 nm in emission. The system was equipped with a Xenon lamp, slit sizes of 5nm (input) and 10 nm (output), and integrated signals over 100 ms. The dataset was imported into Matlab and it was corrected for inner filter effects using the drEEM toolbox [*Murphy and Wuensch*, 2019]. Data were then normalized into Raman units with Rayleigh scattering cut out (both 1st and 2nd order), and Raman scattering interpolated (both 1st and 2nd order).

3.3.6 SPE-DOM extraction and NMR analysis

Solid-phase extraction (SPE) consisted of passing the acidified samples through 1 g of SPE cartridges (PPL, Agilent). Cartridges were soaked overnight in methanol and washed with 150 mL of acidified ultrapure water (AW; pH=2). Samples were extracted at a flow rate of 12 mL/min and rinsed with 100 mL of AW. Following this, cartridges were dried under vacuum and eluted with 20 mL of methanol (2 mL/min).

Extracts were dried completely in a quantitative 2 step-process and redissolved into 600 µL of 50:50 H_2O:D_2O solvent. The spectra were obtained at 298 K on a Bruker DRX 500 spectrometer, equipped with a triple-tuned inverse probe (TXI) operating at 500.13 MHz for ^1H. Water signal suppression was achieved using an SPR-w5-watergate sequence of 125 µs binomial delay, a train of 500, 4 ms calibrated 180 pulses separated by 4 µs delays, corresponding to a 2s saturation loop [*Lam and Simpson*, 2008]. Then the DOM signal was acquired with 128 scans of 16 000 points per

sample: 1.1 s of acquisition time. The total repetition time was 3.2s, while a continuous lock and shimming were applied.

The data was imported into Matlab, and the noise was quantified over the 9 to 11 ppm region. This determined the detection limit as 3 times the standard deviation of the noise, all signals below it were disregarded. The analysis was therefore focused on the signal-rich region of 0.4 to 4.2 ppm, from which the methanol contamination signal [3.27 - 3.30 ppm] was discarded, as were the non-DOM signals of natural silicates [-0.20 - 0.20 ppm] and the residual water peak [4.65 - 4.80 ppm]. From there we integrated the NMR signals by regions as defined in the literature (*Hertkorn et al*. [2013]; *Nebbioso et al*. [2014]; *Pisani et al.* [2015]).

3.3.7 PARAFAC modeling and statistical analysis

PARAFAC analysis from 3 to 7 components was conducted on the EEMs data using the drEEM toolbox. Models were constrained by nonnegativity and a convergence criterion of 10^{-6} was used. The model was validated using residual analysis, split-half analysis, and random initialization as described by *Murphy et al.* [2013]. The best model had 5 components and no outlier samples. Then we extracted and browsed our components through the platform OpenFluor [*Murphy et al.*, 2014].

The statistical differences between harvest days for several parameters were tested through paired t-tests, while differences between size classes were tested by ANOVA. Assumptions were checked using a Shapiro-Wilks normality test of the residuals, and the Bartlett test (homoscedasticity of groups). Non-parametric test of Kruskal-Wallis was used when necessary. Post-hoc analysis was conducted using Tukey's tests (parametric) or Wilcoxon's (non-parametric).

3.4 Results

3.4.1 Growth of phytoplankton assemblages

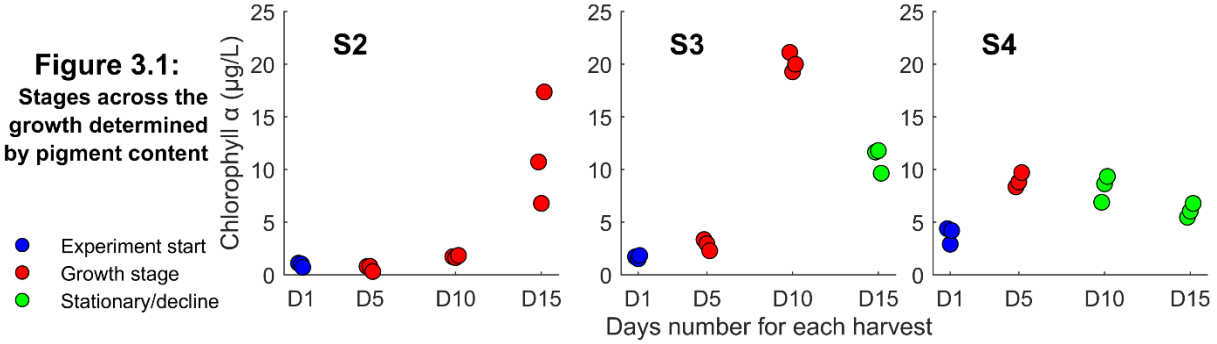


Fig 3.1: *Based on chlorophyll α content, we defined three planktonic development phases (see the text for more details). For more simplicity, the four harvests (days 1, 5, 10, and 15) have been numbered from D1 to D15 and will be named as such in the text.*

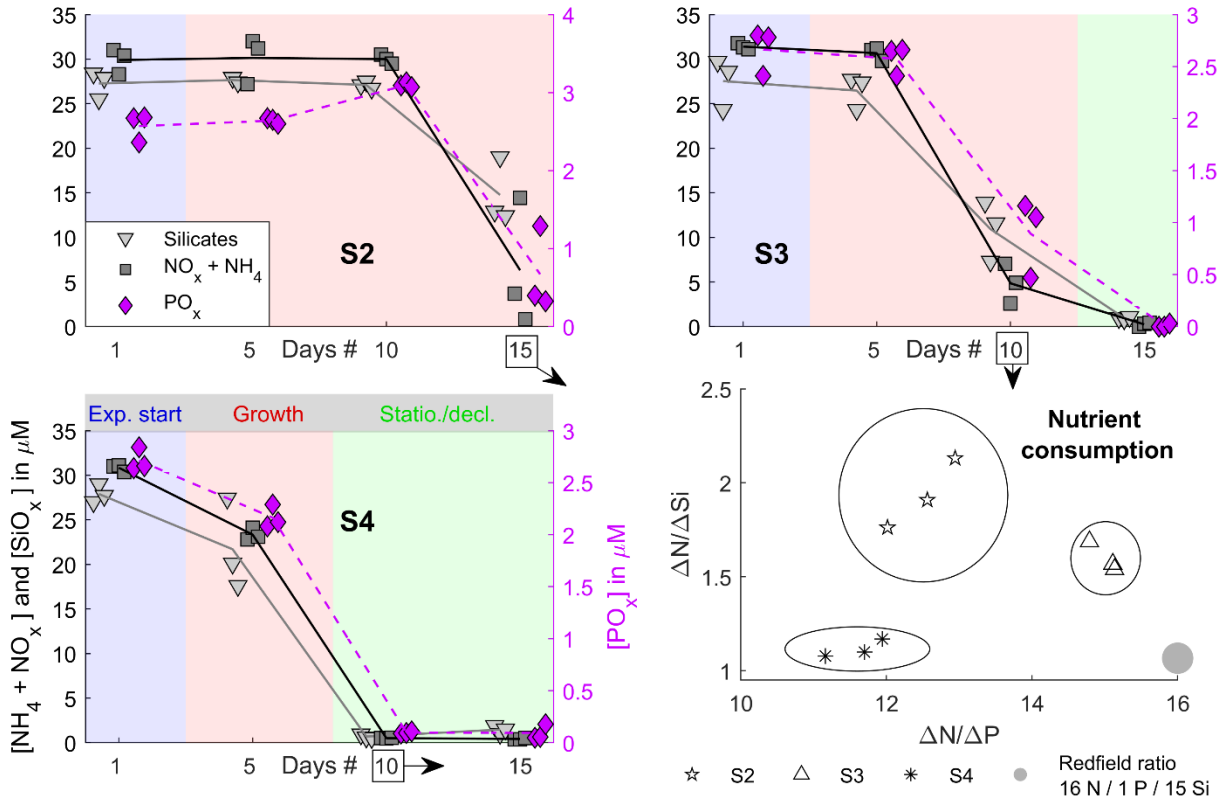


Figure 3.2: Consumption rates of dissolved inorganic nutrients were monitored across the experiment. Growth stages from figure 3.1 are reported using shaded areas. For calculating the rates of nutrient consumption, we selected D15 for S2 and D10 for S3 and S4. Ratios of nutrient consumption are displayed in the bottom left panel. More precisely, the change in N (sum of NO_x and NH_4^+ concentrations) from D1 to the end of growth was compared to the simultaneous depletion of silicate and phosphate. Ellipses correspond to the 95% confidence area based on triplicates' dispersion, and the Redfield ratio is used as a reference point.

Chlorophyll α (Fig. 3.1) demonstrates the growth of algae in all treatments but at different rates. For more simplicity in the text, the four harvest days were termed D1, D5, D10, and D15. Increases of Chl. α were observed as of S2-D10; S3-D5 and S4-D5, indicating that a significant growth started somewhere between those days and the preceding harvest. By contrast, a decline is seen between D15 and the previous harvest for S3 and S4. Chlorophyll α remained stable from D5 to D10 in S4 (i.e., stationary phase), and the inorganic nutrients (SiO_x, PO_x, NH_4^+, and NO_x; Fig. 3.2) became limited suggesting the end of a growth stage after S4-D5 (that is at S4-D10). We restricted "experiment starts" to include only D1 and this placed S2-D5 in the growth stage, despite very small changes in Chl α.

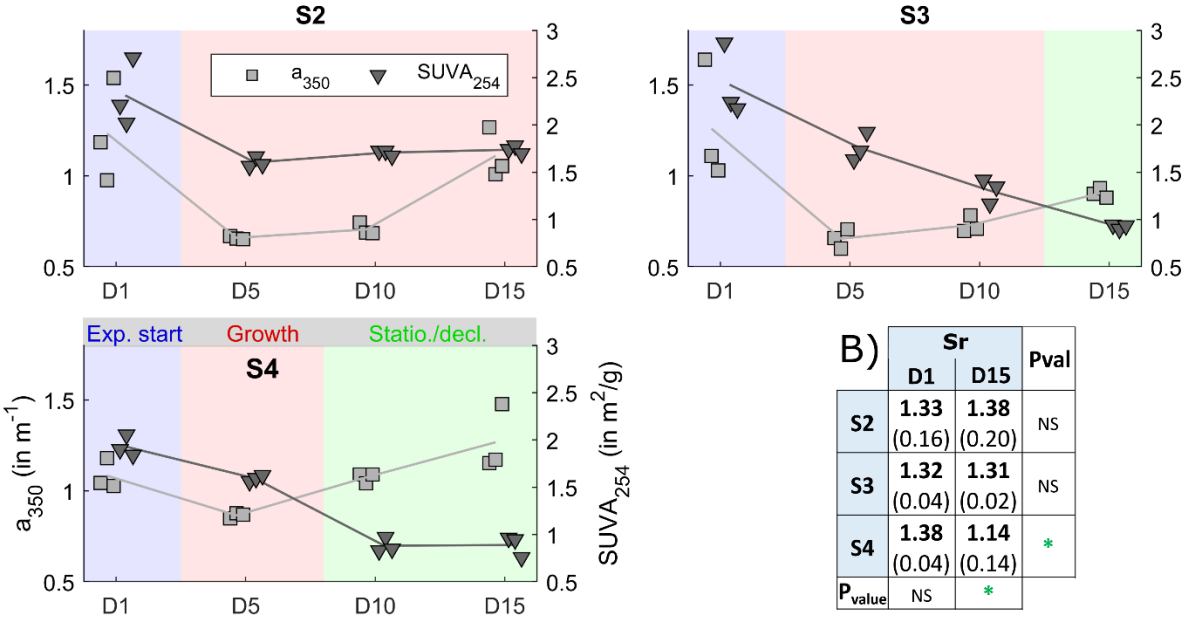
Overall, these patterns support the identification of three growth stages: the experiment start, the growth phase, and the combined stationary/decline phase. Generally, the growth phase corresponds to a period of rapid nutrient consumption and chlorophyll α increase, while the stationary/decline phase was associated with a decline in Chl. α and limited nutrient availability. Importantly, those stages are based on the growth of autotrophs, while heterotrophs (lacking the Chl. α) have likely experienced a different growth pattern.

Inorganic nitrogen (sum of NO_x and NH_4^+) was the first nutrient to be depleted completely in our experiment. For comparing nutrient consumption ratios (Fig. 3.2), we selected S2-D15; S3-D10, and S4-D10 as key time points deeply imprinted by the assemblages' activity. Those time points were the most relevant as the closest (both sides considered) to the real uncaptured tipping point between cell proliferation and cells' stationary phase. It appeared that silicate and phosphate were consumed in different ratios for the 3 communities. While S4 silicate consumption per nitrogen unit (11 N/ 1 P/ 10 Si) was the highest, S3 consumed to a lesser extent its silicate pool (16 N / 1 P/ 10 Si), and S2 even lesser (12 N / 1P/ 6 Si).

The growth of algae in S2 was unexpected as it should only have contained bacteria and little to no chlorophyll (Appendix C). This indicates the failure of the size-fractionation, and observation under a microscope confirmed the presence of diatoms larger than the filter size (0.7 µm) even at D1. Closer inspection revealed that all three treatments appeared to be dominated by centric diatoms, e.g *Chaetoceros sp.*, *Skeletonema sp.*, *Melosira sp.*, *Thalassiossira sp.*, and *Coscinodiscus sp.*, while pennates were only occasionally observed, with no obvious qualitative differences between treatments (Appendix C).

3.4.2 Optical properties of the DOM

Chromophoric DOM (CDOM) properties



B)

	Sr D1	Sr D15	Pval
S2	1.33 (0.16)	1.38 (0.20)	NS
S3	1.32 (0.04)	1.31 (0.02)	NS
S4	1.38 (0.04)	1.14 (0.14)	*
P_{value}	NS	*	

Figure 3.3: Change in CDOM content, a_{350} and $SUVA_{254}$ were monitored across the experiment for each treatment. Growth stages from figure 3.1 are reported using shaded areas. Table B) explores the change in spectral slope ratio (S_{275}/S_{450}) at D1 and D15. The average of triplicates is provided in bold, with the ± value of the 95% confidence interval indicated beneath. The significance of groups' differences was tested by paired t-test and ANOVA, symbols are NS: $p_{value}>0.05$; and *: $p_{value}<0.05$

The change from harvest to harvest of DOM chromophoric properties was interesting, and we aimed in Fig. 3.3 to highlight the time pattern beyond the growth stages. In all treatments, CDOM concentrations dropped from D1 to D5, using the absorption coefficient of $a_{350\ nm}$ as a

proxy, and then increased back to their initial value after growth (D10 and/or D15). Along with this "depletion-recovery" pattern, the aromaticity of the DOM – estimated with $SUVA_{254}$ – underwent diverging trends. While it continuously decreases in S3 and S4, $SUVA_{254}$ remains steady in S2 during the rest of the experiment (D5 to D15).

More generally, the combined information of $SUVA_{254}$ and a_{350} showed the DOM pool at D1 and D5 was similar for all three treatments, with later diverging patterns observed at D15. Moreover, the spectral slope ratio (Sr) was steady throughout the experiment for S2 and S3, while a significant drop occurred for S4 in stationary and decline (Fig. 3.3 B). This significant difference at D15 was not pre-existing at D1, suggesting that the diatoms' growth in treatment S4 resulted in a significantly higher proportion of high molecular weight compounds in the CDOM (p_{value}= 0.03).

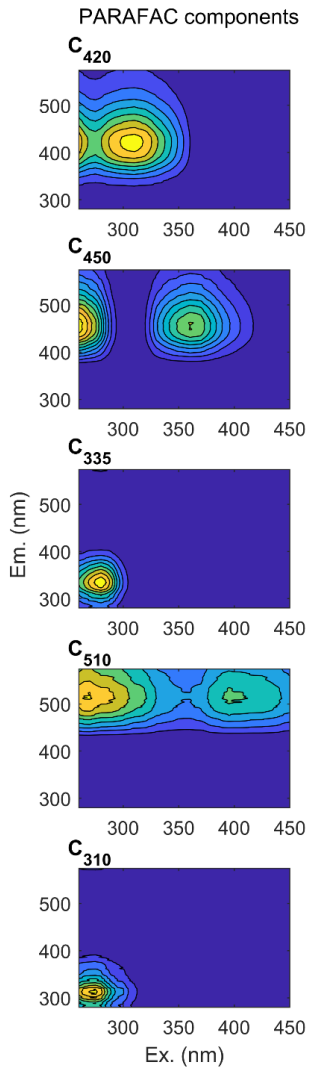
Total DOC content increased through the experiment (Appendix D), but the process lagged behind the growth of primary producers (i.e., Chl. α concentrations). Figure 3.5 A shows that while the chlorophyll α content increased five-fold during the growth period, the DOC content remained stable, and only during the stationary/ decline phase did the concentration of DOC double.

Fluorescent DOM (FDOM): five fluorescent components identified

A PARAFAC model of 5 components (99% of the data variance explained) was found (Fig. 3.4). The components displayed strong similarities (R^2>0.95) with many models previously published and archived in the OpenFluor database [*Murphy et al.*, 2014]. From these, we can ascertain several key properties of our analytes.

C_{420} and C_{450}: We found this component in all samples. In the literature, C_{420} has also been referred to as Peak M $_{310\rightarrow420\ nm}$ and is present in most marine environments and occasionally in freshwater. Many studies have shown its microbial origin through the reprocessing of terrestrial input. Similarly, C_{450} combines the characteristics of peak C $_{360\rightarrow450nm}$ and A $_{270\rightarrow450nm}$, which are

two ubiquitous groups originating from terrestrial-humic structures and persisting in the marine environment [*Coble et al.*, 2014].



Description	Cobles' name	Open Fluor matches	Properties	Reference
310 → 420	Peak M	N=86 ... '' Gao_AF4'' '' Antarctic DOM'' '' Arctic Seawater''	Humic-like fluorescence of microbial origin Present in most marine environment and sometimes in freshwater	(1) (2) (3)
260/360 → 450	Peak C + Peak A	N=11 ... '' Gao_AF4'' ''Kattegat''	Terrestrial-humic origin It persists in the marine environment	(1) (6)
280 → 335	Peak T	N=90 ... '' Gao_AF4'' '' Arctic Seawater'' '' Gueguen_JOIS'' ''AMT20''	Tryptophan-like fluorescence Typically associated to *in-situ*biological production	(1) (3) (4) (5)
270/400 → 510	Peak D	N=61 ... '' Antarctic DOM'' ''Kattegat''	Allochthonous humic-matter, derived from fulvic acids sources Microbially degraded in lakes and estuaries	(2) (6)
270→ 310	Peak B	N=59 ... '' Antarctic_DOM'' ''Kattegat''	Tyrosine-like fluorescence, associated to fresh microbial content Abundant in coastal areas	(2) (6)

Figure 3.4: PARAFAC components' characteristics. The 5 components obtained as a solution were identified through OpenFluor [*Murphy et al.*, 2014] browsing using an R^2=0.95 in both emission and excitation. For each, we provide the heat-map graph associated with the numerical transition. From there Coble's zone names were labeled and confirmed through a literature review. We used the work of ref. (1): model 'Gao_AF4' [*Gao and Guéguen*, 2017] – ref. (2): model 'Antarctic DOM' [*Kida et al.*, 2019] – ref. (3): model 'Arctic Seawater' [*Chen et al.*, 2018] – ref. (4): model 'Gueguen_JOIS' [*DeFrancesco and Guéguen*, 2021] – ref. (5): model 'AMT20' [*Kowalczuk et al.*, 2013] – ref. (6): model 'Kattegat' [*Osburn, and Stedmon*, 2011]

Overall, *Gao and Guéguen* [2017] showed that both C_{420} and C_{450} were composed of a relatively low molecular weight material. Moreover, both are found at all depths in many cold coastal regions [*Chen et al.*, 2018; *Gao and Guéguen*, 2017; *Kida et al*., 2019], where they were negatively correlated to peak $T_{280\rightarrow335nm}$ fluorescence, the latter increasing as seawater warms [*DeFrancesco and Guéguen*, 2021; *Yamashita and Tanoue*, 2003].

C_{510}: This is a microbially labile component typically associated with allochthonous humic substances. C_{510} is quickly degraded in estuaries, yet, bacterial autochthonous production of this component has been recently documented by *Arai et al.* [2017]; and *Fox et al.* [2021], as well as by *Chen et al.* [2018] whose FDOM component matched ours (Fig. 3.4). C_{510} is therefore a component of FDOM that can be both produced and degraded by microbial activity.

C_{335} and C_{310}: The component C_{335} matches the tryptophan-like EEM region (peak $T_{280\rightarrow335nm}$). It is associated with *in-situ* biological production in both marine and freshwater [*Gao and Guéguen*, 2017]. On the other hand, *C_{310}* corresponds to the region in which the amino acid tyrosine fluoresces, sometimes referred to as peak $B_{270\rightarrow310nm}$. This is an abundant component in coastal surface water, associated with fresh microbial production. Warmer zones are associated with a peak B enrichment compared to peak T [*Kowalczuk et al.*, 2013; *Yamashita and Tanoue*, 2003], and trends in the Canadian Arctic suggest increasing spring concentrations with climate change [*DeFrancesco and Guéguen*, 2021].

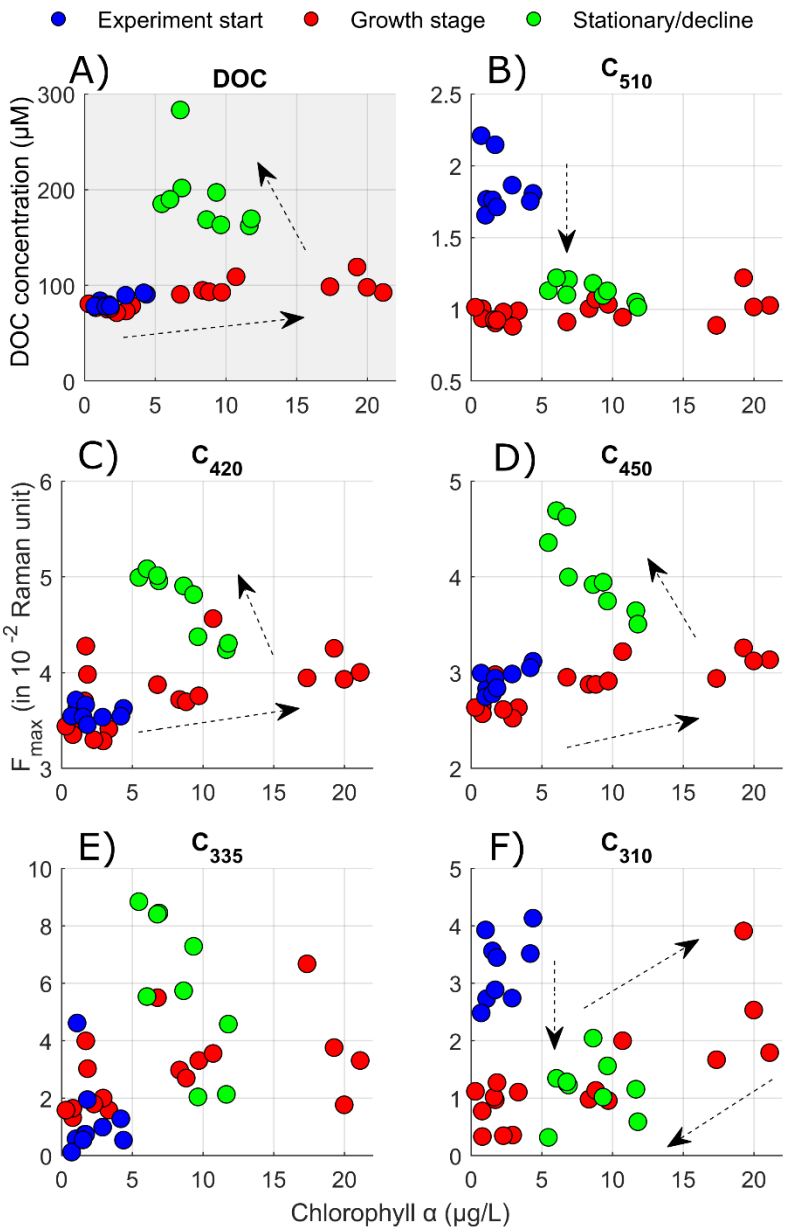


Figure 3.5: FDOM patterns in parallel with DOC production.

Changes in modeled F_{max} are plotted against chlorophyll α content. Across the different components, patterns of consumption, and release are highlighted and compared to the DOC pattern (A).

In our experiment (Fig. 3.5 C), C_{420} concentrations increased only in stationary/decline phases. Similarly (Fig. 3.5 D), concentrations of C_{450} increased by ~50% only after the end of the growth (i.e., when cell lysis is occurring), in a process parallel to the DOC release (Fig. 3.5 A). Thus, our treatments were efficient producers of both kinds of complex chemical structures (visible fluorescence; C_{420} and C_{450}) during the late stages of phototrophic blooms.

In contrast, C_{510} concentration decreased in all treatments within the first few days of the experiment (Fig. 3.5 B). Eventually, a recovery of C_{510} is seen at the end of the growth (~ +15%; $p_{value} = 6 \times 10^{-4}$). Interestingly, the consumption of that fluorescent component occurred while inorganic nutrients were high, while they were low during the production phase (Fig. 3.2).

For the protein-related peaks: levels of C_{335} increased in all cultures across the growth (Fig. 3.5 E). The amount of change varied between treatments with an increase of ~64% in S3 and as much as ~700% in S4. This release was more closely timed with the phytoplankton growth, contrasting with the bulk of DOC exudation (Fig. 3.5 A). Hence, the detectable release of tryptophan-like FDOM occurred in the pool earlier than other DOM proxies (i.e., DOC, C_{420}, C_{450}, C_{510}). Such faster release of C_{335} across growth explains that S2 had the time to accumulate high levels of it (~ +200%), while its levels in those other proxies remained low. Hence, within the selection of DOM proxies, the fast release of C_{335} marks best the active growth phase.

C_{310} was exuded along with phytoplankton growth (Fig. 3.5 F) but with more complex dynamics than C_{335}. At the beginning of the experiment, a significant depletion of C_{310} levels occurred in all treatments, then the levels increased as diatoms growth unfold, before dropping again during the stationary and decline phases. Overall, all assemblages had less C_{310} in the stationary phase relative to the starting conditions. Change in levels of C_{310} appears here as transient and more closely related to cellular stages than C_{335}, as it does not accumulate in the experiment.

3.4.3 NMR content analysis

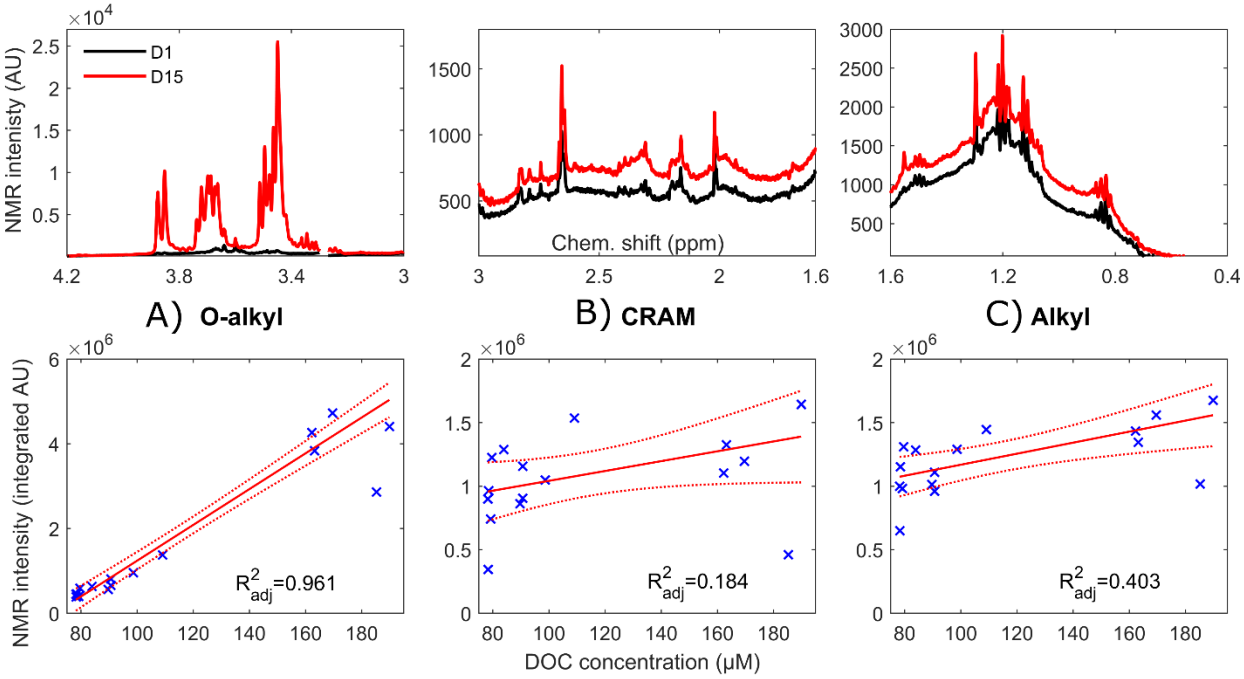


Figure 3.6: ^{1}H-NMR analysis of DOM content following the typical spectra segmentation from the literature for columns A): carbohydrates-related moieties (O-alkyl), B) carboxyl-rich alicyclic molecules (CRAM) and C): saturated aliphatic (alkyl) [*Hertkorn et al.*, 2013; *Nebbioso et al.*, 2014; *Pisani et al.*, 2015]. In the upper panels, averaged shapes on day 1 and day 15 are compared. In lower panels, linear regressions of integrated NMR band intensity to DOC concentration were explored. All models were significant ($p_{value} \leq 0.05$), and the adjusted determination coefficient is indicated. Note the differing y-axes.

Most of the detectable NMR signal was observed from 0.4 to 4.2 ppm (Fig. 3.6). This region includes many moieties which cluster into 3 groups in the literature: alkyl groups, carboxyl-rich alicyclic molecules, and carbohydrates [*Hertkorn et al.*, 2013; *Lam and Simpson*, 2008; *Nebbioso et al.*, 2014]. ^{1}H-NMR can also monitor other functional groups such as "peptides" (4.2-5.2 ppm); "olefinic" moieties (5.2-6.0 ppm); "aromatics" (6.0 - 9.5 ppm) and "amide" structures (7.8-8.4 ppm) [*Pisani et al.*, 2015]. However, in this experiment, these signals were not quantifiable.

<u>*Alkyl and CRAM content:*</u> The "alkyl" band (0.4-1.6 ppm) corresponds to protons at least 4 bonds away from a heteroatom. These carbon-dominated chains can be considered a good

proxy for saturated fatty acids and other saturated alicyclic molecules [*Hertkorn et al*., 2013; *Nebbioso et al*., 2014]. Here (Fig. 3.6 C), the integrated signal of these moieties increased during growth by ~30% (D1 to D15, all treatments combined). However, this NMR signal increase remains lower than the DOC increase (~100%), so while both parameters were strongly correlated ($p_{val} < 0.01$), the bulk of DOC exudation was not from this group. This indicates that the proportion of alkyl moieties in the total DOM pool decreased due to DOC exudation, and no statistical difference in the efficiency of this release was found between treatments (ANOVA).

The band from 1.6 to 3.0 ppm corresponds to aliphatic protons closer to heteroatoms (O and N) than those of the alkyl bands. DOM studies typically associate this group with carboxyl-rich alicyclic molecules (CRAM), a broad cluster of complex structures known to be a major portion of SPE-extractable DOM, as well as acetate derivatives (e.g., $R\text{-}CH_2\text{-}COO^-$) typical of biological material [*Hertkorn et al*., 2013; *Nebbioso et al*., 2014]. Overall, CRAM content increased by ~20% (Fig. 3.6 B) which is 5 times less than the increase in DOC, indicating a non-efficient formation of CRAM while DOC was produced by biological activity.

Carbohydrates production: The "O-alkyl" band (3 – 4.2 ppm) integrates the signals of all protons from single oxygenated units (methoxy derivatives, alcohols, esters, ethers, etc.), and many studies have shown its relation to carbohydrate content in DOM [*Hertkorn et al*., 2013; *Nebbioso et al*., 2014]. Our results show an increase in that signal for all treatments (Fig. 3.6 A; from D1 to D15). This increase correlated strongly with DOC concentration and matched its magnitude of change. It appears therefore that the majority (~85%) of DOC exudates from this plankton community were constituted of "O-alkyl" moieties. As a result, carbohydrates (and associated compounds) passed from constituting ~25% of the DOM pool at the start of the experiment, to ~60% in the stationary phase. This confirms the role of primary producers in the accumulation of those reduced carbon forms.

3.5 Discussion

3.5.1 Contrasting DOM profiles between cultures

Issues in the size-fractionation of the plankton community

The analysis of microbial content showed that diatoms were present in all treatments including S2. The latter was not meant to contain organisms over 0.7 µm, which indicates the failure of the size-fractionation. A perfect separation of plankton groups has been an enduring difficulty in marine studies [*Coveney*, 1978; *Heaney and Jaworski*, 1977; *Schaap et al.*, 2016]. Most likely in our experiment, the multiple steps of liquid transferring allowed the cross-contamination of treatments [*Larsson and Hagström*, 1982; *Porter et al.*, 1993]. Less likely but worth considering, one would note that the 0.7 µm filters had to be changed several times to process the 12 L of water. Each of those changes represented a risk of misplacing the filter (lateral flowthrough via bad sealing), or biomass transfer from the tweezers.

Beyond operational errors for S2, very similar populations of diatoms were found in S3 and S4. Retrospectively, it seems that the experimental design neglected the anatomic difference between cellular length and width. The assumption that bigger diatoms were over 20 µm, is only true considering the primary body axis while the width can be much smaller. For example, the pennates we observed can reach 50 µm across but 12 µm wide. Those diatoms may therefore well have passed vertically through the 20 µm pore size and settled in both S3 and S4. Hence, a potential reassessment of this experiment should consider a rapid observation of the plankton community prior to size-fractionation to identify the most adequate cut-off size for the experiment's purpose.

The difference in treatments' speed of development was also a challenge for data interpretation. Although the size-fractionation failed, it depleted the starting diatoms population by about an order of magnitudes at each filtration step. Therefore, chlorophyll concentrations (proxy of the cell population) took proportionally longer in the lowest density treatment (i.e., S2) than in

the highest density treatment (i.e., S4) to reach the linear range of the Aquapen detector. One drawback of those delays in diatoms' growth period is that DOC production is uneven across treatments (Appendix D). This was particularly true for S2 which did not reach the stationary phase over the course of the experiment. Thus DOC accumulation in S2 was limited, implying fewer changes in several of the proxies we used. In contrast, S4 was in the stationary/ decline phase for over a week and displays DOM markers more strongly impacted by heterotrophic activity.

Chemical properties of the dominant forms of exuded DOM (extracted on PPL)

In the NMR data, most of the increase in signal (~85%) was of carbohydrates-related protons, such that their increase correlated tightly with the change in DOC concentrations across the experiment. The activity of chain-forming algae (primary producers) explains the important production and release of sugars [*Brown*, 1991; *Fogg*, 1966; *Myklestad*, 1995].

Ecological reasons behind diatoms' excretion have been widely explored [*Fogg*, 1966; *Richardson and Cullen*, 1995; *Williams*, 1990]. The conclusions of those studies highlighted cellular homeostasis (e.g., osmotic pressure, mitigation of nutrient limitation, etc.), the beneficial feeding of the phycosphere, and the control of cellular buoyancy control during nitrogen limitation as major drivers of exudation. For instance, species like *Skeletonema costatum* have been shown to store up to 80% of their cellular organic carbon in the form of β-1,3-glucan (high density), which in a normal environment induces cells to sink out of the euphotic zone and below the nutricline [*Granum et al.*, 2002]. There, cells stay in the resting phase until better conditions (e.g., warmer temperature) induce glucan consumption, lowering cellular density and inversing vertical transport. Other forms of carbohydrates, like uronic-acid exopolymers, help in the formation of biofilms by *Melosira sp.* in the epipelagic zone and sea-ice brine channels [*Krembs et al.*, 2011].

Besides the activity of primary producers, our results also show markers of heterotrophic activity [*Xu and Guo*, 2018]. The depletion of CDOM in all treatments and at the early stages of our experiment concords with the results of *Asmala et al.* [2014], where spring CDOM compounds were very labile under nutrient-replete conditions (5% loss in three days of bacterial incubation). Yet here, this consumption was eventually offset by an important CDOM release from the rest of the plankton community, which brought a_{350} levels back to their initial levels.

This depletion-recovery pattern contrast with the $SUVA_{254}$ trends which decrease across the experiment. $SUVA_{254}$ is a proxy associated with aromatics and some related chromophores of the DOM pool, and while this group of compounds is typically produced by bacterial processing [*Kinsey et al.*, 2018], the apparent depletion is mostly induced by the rapid increase of non-aromatic DOC in S3 and S4. Hence, it appears that in our experiment $SUVA_{254}$ behaves as a function of the relative activity of heterotrophs and diatoms (which released most of the DOC). Such DOC production is much smaller in S2 (Appendix D), so its $SUVA_{254}$ remained much steadier.

These patterns in $SUVA_{254}$ could suggest an inefficient production of aromatics in the experiment, and the information from NMR seems to corroborate this with the non-detectable signal over 6 ppm. Aside from aromatics, the CRAM and alkyl group (highly shielded protons) also did not dominate the exudates with only ~5% and ~10% of total NMR-signal change, respectively. Part of the slight increase in NMR's alkyl may come from dissolving cellular fatty acids from diatoms which are typically rich in it [*Brown*, 1991; *Peltomaa et al.*, 2019]. In such a case, a longer experiment (>2 weeks) would likely have observed a more pronounced change in that NMR band, especially as the transfer of fatty acid material into the DOM from cell lysis is a process that takes multiple days [*Engel et al*., 2010; *Parrish*, 1987]. Here, the proportion of CRAM and alkyls in the DOM pool decreased suggesting that the primary producers were not a major

source of these moieties. Indeed, CRAM has long been known to originate from heterotrophic activity [*Kaiser and Benner*, 2008; *Larsson and Hagström*, 1982; *Liu et al.*, 2021].

Finer chemical properties monitored through FDOM

While we observed that the bulk of DOC exudate could be monitored through NMR, this technique remains inaccurate for finer shifts in the DOM pool (i.e., poor signal-to-noise ratio). For instance, changes in aromatic and proteinaceous-related bands could not be reliably quantified, meanwhile strong changes in FDOM composition, some of which are related to those moieties – indicated a rich set of processes [*Ateia et al.*, 2017; *Stubbins et al*., 2014]. Although the FDOM subset of compounds represents only a small fraction of DOC processing, it has been shown as a useful biomarker of multiple processes of the ecosystem [*Amaral et al.*, 2016; *Nieto-Cid et al.*, 2005; *Walker et al.*, 2009; *Yamashita et al.*, 2017].

FDOM production is a bacterial process, and it was shown particularly effective using diatoms' exudates [*Goto et al.*, 2017; *Kuczynska et al.*, 2015; *Romera-Castillo et al.*, 2011]. The processes of consumption or accumulation by the planktonic community are diversified, and in our experiment, four patterns were noted. While C_{420} and C_{450} accumulation followed the same trend as the bulk of DOC exudation, C_{310} preceded it by a few days before dropping again, and C_{510} was essentially degraded before diatoms' growth. Efficiency in C_{335} production was more divergent between treatments (i.e., +65% in S3, +700% in S4), but biological production was nonetheless consistently observed.

Planktonic production of C_{420}, C_{450}, C_{335}, and C_{310} is regularly observed in coastal water during spring bloom conditions [*Nieto-Cid et al.*, 2005; *Yamashita et al.*, 2017; *Yamashita and Tanoue,* 2003]. While C_{335} was associated with tryptophan-like fluorescence in larger peptides, C_{310} was characterized as a marker of smaller peptides and/or tyrosine-like content. Altogether,

they are markers of biological activity related to dissolved proteins, which task is in part to solubilize POC [*Smith et al.*, 1992]. Typically labile, those biomarkers disappear quickly after blooms' end, leaving the longer-lasting imprint of the peak M region (i.e., our C_{420}) and peak C (i.e., C_{450}) [*Jørgensen et al.*, 2014; *Lønborg et al.*, 2010]. Those are particularly ubiquitous in oceanic samples, found at similar levels at all depths and latitudes, and are therefore interpreted as part of the refractory FDOM imprint [*Hansell and Carlson*, 2014; *Jørgensen et al.*, 2011].

In contrast, FDOM components of the C_{510} group were long thought to be allochthonous to coastal areas (i.e., riverine inputs), where it is depleted quickly under the effect of both bio- and photo-degradation [*Coble et al.*, 2014]. Further accumulating evidence showed that plankton production was also a contributor to that pool in the water column [*Arai et al.*, 2017; *Chen et al.*, 2018; *Fox et al.*, 2021; *Kida et al.*, 2019; *Wünsch et al*., 2017]. Concretely, our experiment showed that while C_{510} levels were high at the beginning of the incubation, those levels dropped by half before re-increasing. Such patterns demonstrate the bio-lability and occasional microbial origin of C_{510} at late bloom stages, given that the use of low-UV lamps, multiple layers of glass, and polycarbonate bottles ensured that no photodegradation took place.

Overall, our experiment confirmed that the planktonic community of coastal Newfoundland, and similar regions, can reprocess five distinguishable FDOM components, but each in a unique pattern of consumption and accumulation. In quantitative terms, the growth stage held on average a humic/protein relative fluorescence intensity of ~1.4, while it increased to ~1.8 at stationary and decline phases. Finally, the fluorescence of C_{335} was generally twice as high as that of C_{310}.

The optical properties of S4 were particularly interesting: on one hand, Sr dropped over the course of the experiment, suggesting a shift from the initial DOM pool towards a heavier average molecular weight of CDOM. On the other, C_{335} increased eightfold, indicating significant changes in the protein-related composition of the DOM pool. Those two facts may seem counterintuitive

as, in theory, the physical process of C_{335} fluorescence should generate a steepening of Sr (via an increase in the spectral slope in the UV-B). Yet, it is important to recognize that these two indicators are not mutually exclusive, but rather are driven by different fractions of the DOM pool, which have differing dynamics in this study. Hence, an increase of heavier compounds in S4 may go along with a shift in the peptide-like composition of its FDOM pool.

Another contrasting behavior between treatments was that they displayed diverging inorganic nutrient consumption, and were all different from the theoretical Redfield ratio [*Redfield*, 1958]. This ratio represents an average stoichiometry that varies considerably across taxa and nutrient conditions [*Geider and La Roche*, 2002]; and our results are in agreement with literature values for centric diatoms in xenic cultures like *Chaetoceros spp.*, *Melosira spp.*, *Skeletonema spp.*, and *Thalassiossira spp.* [*Brown*, 1991; *Harrison et al.*, 1977; *Hellebust and Guillard,* 1967; *Spilling*, 2007]. Those studies showed that depending on the species and nutrient conditions their cell metabolism favors different N/Si proportions with values ranging from 1/1 to 2/1. Although we did observe this range of consumption ratios in the inorganic nutrient pool, they differed significantly between treatments, which suggests that the atomic composition of the organic matter (cellular content and/or dissolved) was different. It is important to note that the N/Si is a critical factor controlling the diatoms' growth in the northwest Atlantic [*Romera-Castillo et al.*, 2016]. Hence, the fine processes generating the DOM composition will be likely modulated by changing nutrient conditions.

3.5.2 Time pattern across the experiment

The timing of DOM properties with diatoms growth is another interesting aspect of this study. First, the increases in DOC in all cultures were delayed by a few days compared to the increase in chlorophyll-α content, which operationally monitors diatoms' growth. A key component of DOC exudation is the rise of cellular competition for nutrients, and in our experimental treatments,

nitrogen was found to be limiting in all treatments at the end of the growth phase. At that point, we observed a decline in C_{310} after a sharp increase earlier on. Considering the peptide-like nature of C_{310}, such a process suggests that nitrogen recycling was important in our treatments. *Goto et al.* [2017] observed this release-reabsorption of protein-like fluorochromes in cultures of *Alteromonas sp.* and associated it with cellular uptake. Another study related the quick depletion of protein-like fluorescence in coastal plankton to the aminopeptidase activity (both in cultures of *Skeletonema sp*.; and *Coscinodiscus sp.*) [*Kinsey et al.*, 2018]. Altogether, their results suggest that the biological activity of peptides (release, digestion, and uptake) drives the protein-like presence in the water column.

While most of the changes in FDOM and NMR moieties appeared in nutrient-limited conditions (diatom's stationary and decline phase), the drop of C_{510} occurred in the very first days of the experiment when nutrients were not a stressor. Such a decline of allochthonous FDOM signal followed by a phase of increase was also observed by *Asmala et al.* [2018]. Yet, in another study they found an important production of this FDOM component by the riverine bacterial community (<0.8 μm) feeding over predegraded DOM from the small-end of bacterioplankton (<0.2 μm) [*Asmala et al.*, 2014]. Taken together, these results suggest the successive action of producers and consumers of the moieties appearing as $FDOM_{250(340)\rightarrow 490}$: first fresh material is degraded into non-fluorescing forms while nutrients are high, then the bacterial assemblage adapts to reprocess poorer carbon forms in low nutrients conditions, the end-results of which is fluorescing at the same wavelengths as initially.

Our results support this "consumption-recovery pattern" of complex chemical forms under biological activity. Complex hydrocarbons are generally considered refractory to biodegradation [*Flerus et al.*, 2012; *Lechtenfeld et al.*, 2015]. However, an abundance of inorganic nutrients induces a carbon limitation which can favor bacterial consumption of DOM in coastal

environments [Elovaara et al., 2021; Romera-Castillo et al., 2016; *Shen and Benner*, 2018]. Overall, multiple processes seem to impact the temporal evolution of various DOM characteristics such as DOC concentration, CDOM, and FDOM (in particular C_{510} and C_{310}) highlighting a rich relation to phototrophic activity and secondary metabolism.

3.5.3 Contextualisation of DOM profiles

Contextualization of biological activity in the ecosystem

Monitoring the DOM quality of spring phytoplankton bloom within a laboratory-controlled set-up was an insightful process. While in a real coastal environment part of the biological accumulation of DOM is limited by UV-photodegradation, the use of polycarbonate bottles and PAR-centered low-UV grow lights in our experiment removed this effect to isolate those of biological activity only [*Cory and Kling*, 2018]. Moreover, nutrient availability at the beginning of the experiment was high, so the depletion patterns mimicked in part the natural cycle of spring bloom [Behrenfeld and Boss, 2018]. Grazing zooplankton (larger than 100 µm) were removed before the experiment via size-exclusion, allowing the observation of both diatoms and bacterial metabolism in cooperation, without the confounding effects of sloppy feeding by large grazers. To complete our conclusions on coastal DOM processing in the region, further analysis could explore how salinity and temperature as two important parameters in estuaries [*Singh et al.*, 2019].

The important exudation of carbohydrates was highlighted in the late bloom stage, while CRAM and alkyl moieties did not correlate well with this relatively fresh DOC release. Five fluorescent components were identified and exhibited contrasting patterns, which were associated with the effects of heterotrophic bacteria reprocessing the DOM released by primary producers. The estuarine water initially contained a CDOM pool that heterotrophic bacteria degraded in part before being replenished during diatoms' growth. Generally lower in aromatic proportion, this

fresh material had a specific proteinaceous imprint (C_{335} vs C_{310}), and in at least one instance, a higher proportion of heavy CDOM.

This description fits previous descriptions of phytoplankton blooms in other regions across the cold-temperate ecosystem of the Northern Hemisphere [*Carr et al.*, 2019; *Hansman et al.*, 2015; *Heal et al.*, 2020; *Ittekkot et al.*, 1981]. Among the most recent research, *Merder et al.* [2021] subdivided 1392 molecular formulae (identified via mass spectrometry) from coastal DOM into 11 clusters, five of which directly correlated to the abundance of assessed plankton species. Those compounds had a low aromaticity index, little oxidation (O/C), and a higher proportion of peptide-like molecular formulae. Cluster abundances across the year were spiky, with multiple 1-2 week hot-spot occurring repeatedly across the time series. Altogether those parameters distinguished those aliphatic molecules from the other terrestrial and "refractory marine DOM" groups. Many elements of this description match ours, and just as we noted in our experiment the dominance of exudates in the NMR's 3-4 ppm region (O-alkyl), the MS atomic composition they found (plankton-related DOM) was at ~70% CHO forms.

Overall, the experiment demonstrated the production of very labile carbon forms. If this DOM quality mirrors the nutritional values at the base of the food web (e.g., phytoplankton-related carbohydrates and lipids), then such planktonic assemblage should fuel efficiently further levels of consumption. Effectively, beyond bacterial consumption, the low energetic requirements of those chemicals enable as well high growth efficiency of grazers, filter feeders (e.g., krill), and subsequent animals [*Pershing and Stamieszkin*, 2019; *Zavorka et al.*, 2021]. Although our experiment did not assess if the DOM lability marked plankton feeding quality, many studies suggest it could [*Engel et al.*, 2010; *Granum et al.*, 2002; *Williams*, 1990]. Hence monitoring the signature of DOM offers us a glimpse of shifts in resource quality in the environment.

How could climate change shift the DOM profiles?

Shifts in biological sources and processes of DOM under changing environmental conditions remains poorly forecasted by the current biogeochemical paradigm [*Anderson et al.*, 2015; *Thornton,* 2014]. From oceanographic records, one could hypothesize that processes occurring in warmer zones (i.e., Northeastern US coast) may soon take place in coastal Newfoundland [*Greene et al.*, 2008]. Although caution should be taken in such comparison, it appears that heterotrophic activity should be boosted by warmer conditions and that diatoms' gross primary production may become N limited (N/Si ratio) [*Balaguru et al.,* 2018; *Rivkin et al.,* 1996]. As a result, CDOM could tend to be lower in concentration (i.e., a_{350}), with a compositional shift with enriched aromatics ($SUVA_{254}$). Further evidence from oceanic transects (pelagic) suggests that proteinaceous FDOM components could increase, especially C_{310}, while marine-humic components (C_{420} and C_{450}) could comparatively decrease [*Alonso-Sáez et al.*, 2012; *Garcia-Corral et al.*, 2017; *Kowalczuk et al.,* 2013; *Yamashita et al.*, 2017].

Although C_{510} terrestrial inputs will likely increase in estuaries, the microbial carbon processing in coastal ecosystems will also likely adapt to warmer conditions beyond riverine influence. In fact, *Dawson et al.* [2020] found that warmer and fresher water in the coastal Arctic Ocean should induce higher content of tyrosine, tryptophan, and certain Krebs' cycle intermediates, along with a decrease in carbon-rich metabolites used in glycolysis. Thus, monitoring those shifts may therefore reveal the shift of the coastal Newfoundland DOM fingerprint towards southern properties.

Conclusion

This experiment monitored the biological reprocessing of estuarine natural DOM in three diatom assemblages. Several points of the growth curve were covered, and it appeared that the bulk of DOC exudation was carbohydrate-related and released in late bloom stages. Five fluorescent components were identified and produced contrasting patterns, some in parallel to DOC exudation and others earlier in the growth phase. While the heterotrophic consumption of CDOM and the FDOM component C_{510} was followed by a production later in the experiment; the aromatic proportion of the pool generally decreased across the experiment. Among the proteinaceous-related FDOM components, C_{310} was shown to disappear more quickly than C_{335} from the water column. The properties of this fresh material compared to *in-situ* studies suggests the sensitivity of the DOM quality to the N/Si ratio as climate change unfolds.

3.6 Reference

Alonso-Sáez, L., O. Sánchez, and J. M. Gasol (2012), Bacterial uptake of low molecular weight organics in the subtropical Atlantic: Are major phylogenetic groups functionally different?, *Limnology and Oceanography*, *57*(3), 798-808, doi:https://doi.org/10.4319/lo.2012.57.3.0798.

Amaral, V., D. Graeber, D. Calliari, and C. Alonso (2016), Strong linkages between DOM optical properties and main clades of aquatic bacteria, *Limnology and Oceanography*, *61*(3), 906-918, doi:10.1002/lno.10258.

Aminot, A., and F. Rey (2002), Standard procedure for the determination of chlorophyll a by spectroscopic methods, *Rep.*, 25 pp, International Council for the Exploration of the Sea (ICES), Copenhagen, Denmark, doi: https://doi.org/10.17895/ices.pub.5080.

Anderson, T. R., J. R. Christian, and K. J. Flynn (2015), Chapter 15 - Modeling DOM biogeochemistry, in *Biogeochemistry of marine dissolved organic matter (second edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 635-667, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00015-7.

Arai, K., S. Wada, K. Shimotori, Y. Omori, and T. Hama (2017), Production and degradation of fluorescent dissolved organic matter derived from bacteria, *Journal of Oceanography*, *74*(1), 39-52, doi:https://doi.org/10.1007/s10872-017-0436-y.

Asmala, R. Autio, H. Kaartokallio, C. A. Stedmon, and D. N. Thomas (2014), Processing of humic-rich riverine dissolved organic matter by estuarine bacteria: effects of predegradation and inorganic nutrients, *Aquatic Sciences*, *76*(3), 451-463, doi:10.1007/s00027-014-0346-7.

Asmala, L. Haraguchi, H. H. Jakobsen, P. Massicotte, and J. Carstensen (2018), Nutrient availability as major driver of phytoplankton-derived dissolved organic matter transformation in coastal environment, *Biogeochemistry*, *137*(1), 93-104, doi:10.1007/s10533-017-0403-0.

Ateia, M., J. Ran, M. Fujii, and C. Yoshimura (2017), The relationship between molecular composition and fluorescence properties of humic substances, *International Journal of Environmental Science and Technology*, *14*(4), 867-880, doi:10.1007/s13762-016-1214-x.

Baker, A. L. (2012), Phycokey -- an image based key to Algae (PS Protista), Cyanobacteria, and other aquatic objects, [online database], edited, Center for Freshwater Biology, University of New Hampshire, URL: http://cfb.unh.edu/phycokey/phycokey.htm.

Balaguru, K., S. C. Doney, L. Bianucci, P. J. Rasch, L. R. Leung, J. H. Yoon, and I. D. Lima (2018), Linking deep convection and phytoplankton blooms in the northern Labrador Sea in a changing climate, *PLoS One*, *13*(1), 17 p., doi:10.1371/journal.pone.0191509.

Behrenfeld, M. J., and E. S. Boss (2018), Student's tutorial on bloom hypotheses in the context of phytoplankton annual cycles, *Global Change Biology*, *24*(1), 55-77, doi:10.1111/gcb.13858.

Broughton, D. (1958-2020), CPR Survey Data Catalog., [online database- version 2020- consulted on 02/2022], edited, Marine Biological Association of the UK (MBA), Plymouth, UK, URL: https://www.cprsurvey.org/data/map-data/.

Brown, M. R. (1991), The amino-acid and sugar composition of 16 species of microalgae used in mariculture, *Journal of Experimental Marine Biology and Ecology*, *145*(1), 79-99, doi:https://doi.org/10.1016/0022-0981(91)90007-J.

Buck, K. R., F. P. Chavez, and L. Campbell (1996), Basin-wide distributions of living carbon components and the inverted trophic pyramid of the central gyre of the North Atlantic Ocean, summer 1993, *Aquatic Microbial Ecology*, *10*(3), 283-298, doi:10.3354/ame010283.

Carlson, C. A., and D. A. Hansell (2015), Chapter 3 - DOM sources, sinks, reactivity, and budgets, in *Biogeochemistry of marine dissolved organic matter (second edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 65-126, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00003-0.

Carr, N., C. E. Davis, S. Blackbird, L. R. Daniels, C. Preece, M. Woodward, and C. Mahaffey (2019), Seasonal and spatial variability in the optical characteristics of DOM in a temperate shelf sea, *Progress in Oceanography*, *177*, 101929, doi:https://doi.org/10.1016/j.pocean.2018.02.025.

Catalá, T. S., S. Shorte, and T. Dittmar (2021), Marine dissolved organic matter: a vast and unexplored molecular space, *Applied Microbiology and Biotechnology*, *105*(19), 7225-7239, doi:10.1007/s00253-021-11489-3.

Chen, M., J. Jung, Y. K. Lee, and J. Hur (2018), Surface accumulation of low molecular weight dissolved organic matter in surface waters and horizontal off-shelf spreading of nutrients and humic-like fluorescence in the Chukchi Sea of the Arctic Ocean, *Science of The Total Environment*, *639*, 624-632, doi:https://doi.org/10.1016/j.scitotenv.2018.05.205.

Coble, P. G., J. Lead, A. Baker, D. M. Reynolds, and R. Spencer (2014), *Aquatic organic matter fluorescence*, 375 pp., Cambridge University Press, Cambridge, doi:10.1017/CBO9781139045452.

Cory, R. M., and G. W. Kling (2018), Interactions between sunlight and microorganisms influence dissolved organic matter degradation along the aquatic continuum, *Limnology and Oceanography Letters*, *3*(3), 102-116, doi:https://doi.org/10.1002/lol2.10060.

Coutinho, F., P. Meirelles, A. P. Moreira, R. Paranhos, B. Dutilh, and F. Thompson (2015), Niche distribution and influence of environmental parameters in marine microbial communities: a systematic review, *PeerJ*, *3*, 19 p., doi:10.7717/peerj.1008.

Coveney, M. F. (1978), Separation of algae and bacteria in lake water by size fractionation, *SIL Proceedings, 1922-2010*, *20*(2), 1264-1269, doi:10.1080/03680770.1977.11896684.

Daniels, C., A. Poulton, M. Esposito, M. Paulsen, R. Bellerby, A. Martin, and M. John (2015), Phytoplankton dynamics in contrasting early stage North Atlantic spring blooms: Composition, succession, and potential drivers, *Biogeosciences*, *12*(8), 2395-2409, doi:10.5194/bg-12-2395-2015.

Dawson, H. M., K. R. Heal, A. K. Boysen, L. T. Carlson, A. E. Ingalls, and J. N. Young (2020), Potential of temperature- and salinity-driven shifts in diatom compatible solute concentrations to impact biogeochemical cycling within sea ice, *Elementa: Science of the Anthropocene*, *8*, doi:10.1525/elementa.421.

DeFrancesco, C., and C. Guéguen (2021), Long-term trends in dissolved organic matter composition and its relation to sea ice in the Canada basin, Arctic Ocean (2007–2017), *Journal of Geophysical Research: Oceans*, *126*(2), 14 p., doi:https://doi.org/10.1029/2020JC016578.

Elovaara, S., E. Eronen-Rasimus, E. Asmala, T. Tamelander, and H. Kaartokallio (2021), Contrasting patterns of carbon cycling and dissolved organic matter processing in two phytoplankton–bacteria communities, *Biogeosciences*, *18*(24), 6589-6616, doi:10.5194/bg-18-6589-2021.

Engel, A., N. Händel, J. Wohlers, M. Lunau, H.-P. Grossart, U. Sommer, and U. Riebesell (2010), Effects of sea surface warming on the production and composition of dissolved organic matter during phytoplankton blooms: results from a mesocosm study, *Journal of Plankton Research*, *33*(3), 357-372, doi:10.1093/plankt/fbq122.

Figueroa, D. (2016), Bacterioplankton in the Baltic Sea -Influence of allochthonous organic matter and salinity, [Thesis], 23 pp, Umeå University, Sweden, ISBN 978-91-7601-412-7.

Flerus, R., O. Lechtenfeld, B. Koch, S. McCallister, P. Schmitt-Kopplin, R. Benner, K. Kaiser, and G. Kattner (2012), A molecular perspective on the ageing of marine dissolved organic matter, *Biogeosciences*, *9*(6), 1935–1955, doi:10.5194/bg-9-1935-2012.

Fogg, G. E. (1966), The extracellular products of algae, *Oceanogr. Mar. Biol. Ann. Rev.*, *4*, 195-212.

Fox, B., R. Thorn, and D. Reynolds (2021), Laboratory in-situ production of autochthonous and allochthonous fluorescent organic matter by freshwater bacteria, *Microorganisms*, *9*(8), 1623, doi:10.3390/microorganisms9081623.

Gao, Z., and C. Guéguen (2017), Size distribution of absorbing and fluorescing DOM in Beaufort Sea, Canada Basin, *Deep Sea Research Part I: Oceanographic Research Papers*, *121*, 30-37, doi:https://doi.org/10.1016/j.dsr.2016.12.014.

Garcia-Corral, L. S., et al. (2017), Temperature dependence of plankton community metabolism in the subtropical and tropical oceans, *Global Biogeochemical Cycles*, *31*(7), 1141-1154, doi:https://doi.org/10.1002/2017GB005629.

Geider, R., and J. La Roche (2002), Redfield revisited: variability of C:N:P in marine microalgae and its biochemical basis, *European Journal of Phycology*, *37*(1), 1-17, doi:10.1017/S0967026201003456.

Goto, S., Y. Tada, K. Suzuki, and Y. Yamashita (2017), Production and reutilization of fluorescent dissolved organic matter by a marine bacterial strain, Alteromonas macleodii, *Front Microbiol*, *28*(8), 507, doi:10.3389/fmicb.2017.00507.

Granum, E., S. Kirkvold, and S. Myklestad (2002), Cellular and extracellular production of carbohydrates and amino acids by the marine diatom Skeletonema costatum: Diel variations and effects of N depletion, *Marine Ecology-progress Series*, *242*, 83-94, doi:10.3354/meps242083.

Greene, C. H., A. J. Pershing, T. M. Cronin, and N. Ceci (2008), Arctic climatechange and its impacts on the ecology of the North Atlantic, *Ecology*, *89*(11 S), 24-38, doi:10.1890/07-0550.1.

Hansell, D. A., and C. A. Carlson (2014), *Biogeochemistry of marine dissolved organic matter*, 698 pp., Elsevier Science & Technology, San Diego, US, ISBN 9780124071537.

Hansman, R. L., T. Dittmar, and G. J. Herndl (2015), Conservation of dissolved organic matter molecular composition during mixing of the deep water masses of the northeast Atlantic Ocean, *Marine Chemistry*, *177*(2), 288-297, doi:https://doi.org/10.1016/j.marchem.2015.06.001.

Harrison, P. J., H. L. Conway, R. W. Holmes, and C. O. Davis (1977), Marine diatoms grown in chemostats under silicate or ammonium limitation. III. Cellular chemical composition and morphology of Chaetoceros debilis, Skeletonema costatum, and Thalassiosira gravida, *Marine Biology*, *43*(1), 19-31, doi:10.1007/BF00392568.

Head, E. J., and D. D. Sameoto (2007), Inter-decadal variability in zooplankton and phytoplankton abundance on the Newfoundland and Scotian shelves, *Deep Sea Research Part II: Topical Studies in Oceanography*, *54*(23), 2686-2701, doi:https://doi.org/10.1016/j.dsr2.2007.08.003.

Heal, K. R., B. Durham, A. K. Boysen, L. T. Carlson, W. Qin, F. Ribalet, A. E. White, R. M. Bundy, E. V. Armbrust, and A. E. Ingalls (2020), Marine community metabolomes carry fingerprints of phytoplankton community composition, *bioRxiv*, *6*(3), 19 p., doi:10.1101/2020.12.22.424086.

Heaney, S. I., and G. H. Jaworski (1977), A simple separation technique for purifying micro-algae, *British Phycological Journal*, *12*(2), 171-174, doi:10.1080/00071617700650191.

Hedges, J. (2002), Chapter 1- Why dissolved organics matter, in *Biogeochemistry of Marine Dissolved Organic Matter, 1st edition*, edited by D. A. Hansell and C. A. Carlson, pp. 1–33, Academic Press, San Diego, doi:10.1016/B978-012323841-2/50003-8.

Hellebust, J. A., and R. R. Guillard (1967), Uptake specificity for organic substrates by the marine Diatom Melosira Nummuloides, *J Phycol*, *3*(3), 132-136, doi:10.1111/j.1529-8817.1967.tb04646.x.

Helms, J. R., A. Stubbins, J. D. Ritchie, E. C. Minor, D. J. Kieber, and K. Mopper (2008), Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter, *Limnology and Oceanography*, *53*(3), 955-969, doi:10.4319/lo.2008.53.3.0955.

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*(3), 1583-1624, doi:10.5194/bg-10-1583-2013.

Hinder, S., G. Hays, M. Edwards, E. Roberts, A. Walne, and M. Gravenor (2012), Changes in marine dinoflagellate and diatom abundance under climate change, *Nature Climate Change*, *2*, 271-275, doi:10.1038/nclimate1388.

Ittekkot, V., U. Brockmann, W. Michaelis, and E. T. Degens (1981), Dissolved free and combined carbohydrates during a phytoplankton bloom in the northern North Sea, *Marine Ecology Progress Series*, *4*(3), 299-305, URL: https://www.jstor.org/stable/24813224.

Jeffrey, S. W., and G. F. Humphrey (1975), New spectrophotometric equations for determining chlorophylls a, b, c1 and c2 in higher plants, algae and natural phytoplankton, *Biochemie und Physiologie der Pflanzen*, *167*(2), 191-194, doi:https://doi.org/10.1016/S0015-3796(17)30778-3.

Jiao, N., et al. (2010), Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean, *Nature Reviews Microbiology*, *8*(8), 593-599, doi:10.1038/nrmicro2386.

Jørgensen, L., C. A. Stedmon, M. A. Granskog, and M. Middelboe (2014), Tracing the long-term microbial production of recalcitrant fluorescent dissolved organic matter in seawater, *Geophysical Research Letters*, *41*(7), 2481-2488, doi:https://doi.org/10.1002/2014GL059428.

Jørgensen, L., C. A. Stedmon, T. Kragh, S. Markager, M. Middelboe, and M. Søndergaard (2011), Global trends in the fluorescence characteristics and distribution of marine dissolved organic matter, *Marine Chemistry*, *126*(1), 139-148, doi:https://doi.org/10.1016/j.marchem.2011.05.002.

Kaiser, K., and R. Benner (2008), Major bacterial contribution to the ocean reservoir of detrital organic carbon and nitrogen, *Limnology and Oceanography*, *53*(1), 99-112, doi:https://doi.org/10.4319/lo.2008.53.1.0099.

Kida, M., T. Kojima, Y. Tanabe, K. Hayashi, S. Kudoh, N. Maie, and N. Fujitake (2019), Origin, distributions, and environmental significance of ubiquitous humic-like fluorophores in Antarctic lakes and streams, *Water Research*, *163*(114901), 11 p., doi:https://doi.org/10.1016/j.watres.2019.114901.

Kinsey, J., G. Corradino, K. Ziervogel, A. Schnetzer, and C. Osburn (2018), Formation of chromophoric dissolved organic matter by bacterial degradation of phytoplankton-derived aggregates, *Frontiers in Marine Science*, *4*(430), 16 p., doi:10.3389/fmars.2017.00430.

Kowalczuk, P., G. H. Tilstone, M. Zabłocka, R. Röttgers, and R. Thomas (2013), Composition of dissolved organic matter along an Atlantic Meridional Transect from fluorescence spectroscopy and Parallel Factor Analysis, *Marine Chemistry*, *157*, 170-184, doi:https://doi.org/10.1016/j.marchem.2013.10.004.

Krembs, C., H. Eicken, and J. W. Deming (2011), Exopolymer alteration of physical properties of sea ice and implications for ice habitability and biogeochemistry in a warmer Arctic, *Proc Natl Acad Sci U S A*, *108*(9), 3653-3658, doi:10.1073/pnas.1100701108.

Kuczynska, P., M. Jemiola-Rzeminska, and K. Strzalka (2015), Photosynthetic pigments in diatoms, *Marine drugs*, *13*(9), 5847-5881, doi:10.3390/md13095847.

LACHAT intruments (2021), Methods list for automated ion analyzers- revision 17, in *Flow Injection Analysis • Ion Chromatography*, edited, Hach company, URL: https://my.hach.com/family-downloads.jsa?productCategoryId=59429623939.

Lam, B., and A. J. Simpson (2008), Direct 1H NMR spectroscopy of dissolved organic matter in natural waters, *Analyst*, *133*(2), 263-269, doi:10.1039/B713457F.

Larsson, U., and Å. Hagström (1982), Fractionated phytoplankton primary production, exudate release and bacterial production in a Baltic eutrophication gradient, *Marine Biology*, *67*(1), 57-70, doi:10.1007/BF00397095.

Lechtenfeld, O. J., N. Hertkorn, Y. Shen, M. Witt, and R. Benner (2015), Marine sequestration of carbon in bacterial metabolites, *Nat Commun*, *6*(1), 6711, doi:10.1038/ncomms7711.

Lewin, J. C., and R. A. Lewin (1960), Auxotrophy and heterotrophy in marine littoral diatoms, *Can J Microbiol*, *6*(2), 127-134, doi:10.1139/m60-015.

Liu, Y., J. Kan, C. He, Q. Shi, Y. X. Liu, Z. C. Fan, and J. Sun (2021), Epiphytic bacteria are essential for the production and transformation of algae-derived carboxyl-rich alicyclic molecule (CRAM)-like DOM, *Microbiol Spectr*, *9*(2), 16 p., doi:10.1128/Spectrum.01531-21.

Lønborg, C., X. A. Álvarez-Salgado, K. Davidson, S. Martínez-García, and E. Teira (2010), Assessing the microbial bioavailability and degradation rate constants of dissolved organic matter by fluorescence spectroscopy in the coastal upwelling system of the Ría de Vigo, *Marine Chemistry*, *119*(1), 121-129, doi:https://doi.org/10.1016/j.marchem.2010.02.001.

Lynch, L. M., N. A. Sutfin, T. S. Fegel, C. M. Boot, T. P. Covino, and M. D. Wallenstein (2019), River channel connectivity shifts metabolite composition and dissolved organic matter chemistry, *Nature Communications*, *10*(1), 459, doi:10.1038/s41467-019-08406-8.

Marie, D., F. Partensky, D. Vaulot, and C. Brussaard (2001), Enumeration of phytoplankton, bacteria, and viruses in marine samples (chap 11.11-supplement), in *Current Protocols in Cytometry*, edited, pp. 9-10, doi:https://doi.org/10.1002/0471142956.CY1111S10.

Merder, J., H. Röder, T. Dittmar, U. Feudel, J. Freund, G. Gerdts, A. Kraberg, and J. Niggemann (2021), Dissolved organic compounds with synchronous dynamics share chemical properties and origin, *Limnology and Oceanography*, *66*(11), 4001-4016, doi:10.1002/lno.11938.

Murphy, K. R., C. A. Stedmon, D. Graeber, and R. Bro (2013), Fluorescence spectroscopy and multi-way techniques: PARAFAC, *Analytical Methods*, *5*(23), 6557-6566, doi:10.1039/C3AY41160E.

Murphy, K. R., C. A. Stedmon, P. Wenig, and R. Bro (2014), OpenFluor- an online spectral library of auto-fluorescence by organic compounds in the environment, *Analytical Methods*, *6*(3), 658-661, doi:10.1039/C3AY41935E.

Murphy, K. R., and U. Wuensch (2019), drEEM toolbox (C) 0.5.1, [computer program], edited, Chalmers University of Technology, URL: http://models.life.ku.dk/drEEM.

Myklestad, S. M. (1995), Release of extracellular products by phytoplankton with special emphasis on polysaccharides, *Science of The Total Environment*, *165*(1), 155-164, doi:https://doi.org/10.1016/0048-9697(95)04549-G.

Nebbioso, A., P. Mazzei, and D. Savy (2014), Reduced complexity of multidimensional and diffusion NMR spectra of soil humic fractions as simplified by Humeomics, *Chemical and Biological Technologies in Agriculture*, *1*(1), 24, doi:10.1186/s40538-014-0024-y.

Nieto-Cid, M., X. A. Álvarez-Salgado, J. Gago, and F. F. Pérez (2005), DOM fluorescence, a tracer for biogeochemical processes in a coastal upwelling system (NW Iberian Peninsula), *Marine Ecology Progress Series*, *297*, 33-50, URL: http://www.jstor.org/stable/24868645.

Parrish, C. C. (1987), Time series of particulate and dissolved lipid classes during spring phytoplankton blooms in Bedford Basin, a marine inlet, *Marine Ecology Progress Series*, *35*(1-2), 129-139, URL: https://www.jstor.org/stable/24825017.

Peltomaa, E., H. Hällfors, and S. J. Taipale (2019), Comparison of diatoms and dinoflagellates from different habitats as sources of PUFAs, *Mar Drugs*, *17*(4), doi:10.3390/md17040233.

Pershing, A. J., and K. Stamieszkin (2019), The north Atlantic ecosystem, from plankton to whales, *Ann Rev Mar Sci*, *12*(1), 339-359, doi:10.1146/annurev-marine-010419-010752.

Pisani, O., S. D. Frey, A. J. Simpson, and M. J. Simpson (2015), Soil warming and nitrogen deposition alter soil organic matter composition at the molecular-level, *Biogeochemistry*, *123*(3), 391-409, doi:10.1007/s10533-015-0073-8.

Porter, J., C. Edwards, J. A. Morgan, and R. W. Pickup (1993), Rapid, automated separation of specific bacteria from lake water and sewage by flow cytometry and cell sorting, *Applied and environmental microbiology*, *59*(10), 3327-3333, doi:10.1128/aem.59.10.3327-3333.1993.

Redfield, A. C. (1958), The biological control of chemical factors in the environment, *Science progress*, *11*, 150-170, PMID: 24545739.

Richardson, T., and J. Cullen (1995), Changes in buoyancy and chemical composition during growth of a coastal marine diatom: Ecological and biogeochemical consequences, *Marine Ecology Progress Series*, *128*(1), 77-90, doi:10.3354/meps128077.

Rivkin, R. B., M. R. Anderson, and C. Lajzerowicz (1996), Microbial processes in cold oceans. I. Relationship between temperature and bacterial growth rate, *Aquatic Microbial Ecology*, *10*(3), 243-254, doi:10.3354/ame010243.

Romera-Castillo, C., R. T. Letscher, and D. A. Hansell (2016), New nutrients exert fundamental control on dissolved organic carbon accumulation in the surface Atlantic Ocean, *Proceedings of the National Academy of Sciences*, *113*(38), 10497-10502, doi:10.1073/pnas.1605344113.

Romera-Castillo, C., H. Sarmento, X. A. Alvarez-Salgado, J. M. Gasol, and C. Marrasé (2011), Net production and consumption of fluorescent colored dissolved organic matter by natural bacterial assemblages growing on marine phytoplankton exudates, *Appl Environ Microbiol*, *77*(21), 7490-7498, doi:10.1128/aem.00200-11.

Schaap, A., J. Dumon, and J. d. Toonder (2016), Sorting algal cells by morphology in spiral microchannels using inertial microfluidics, *Microfluidics and Nanofluidics*, *20*(9), 125, doi:10.1007/s10404-016-1787-1.

Shen, Y., and R. Benner (2018), Mixing it up in the ocean carbon cycle and the removal of refractory dissolved organic carbon, *Scientific Reports*, *8*(1), 2542, doi:10.1038/s41598-018-20857-5.

Shibl, A. A., et al. (2020), Diatom modulation of select bacteria through use of two unique secondary metabolites, *Proceedings of the National Academy of Sciences*, *117*(44), 27445-27455, doi:10.1073/pnas.2012088117.

Singh, S., P. Dash, M. S. Sankar, S. Silwal, Y. Lu, P. Shang, and R. J. Moorhead (2019), Hydrological and biogeochemical controls of seasonality in dissolved organic matter delivery to a blackwater estuary, *Estuaries and Coasts*, *42*(2), 439-454, doi:10.1007/s12237-018-0473-9.

Smith, D. C., M. Simon, A. L. Alldredge, and F. Azam (1992), Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution, *Nature*, *359*(6391), 139-142, doi:10.1038/359139a0.

Spice, W. M., and J. P. Ackers (1992), The effect of axenic versus xenic culture conditions on the total and secreted proteolytic activity of Entamoeba histolytica strains, *Arch Med Res*, *23*(2), 91-93, PMID: 1340330.

Spilling, K. (2007), On the ecology of cold-water phytoplankton in the Baltic Sea, in *Scientific Report No. 31*, edited, pp. 1-59, W. & A. de Nottbeck Foundation, ISBN 978-952-10-3626-2.

Stubbins, A., J. F. Lapierre, M. Berggren, Y. T. Prairie, T. Dittmar, and P. A. del Giorgio (2014), What's in an EEM? Molecular signatures associated with dissolved organic fluorescence in boreal Canada, *Environ Sci Technol*, *48*(18), 10598-10606, doi:10.1021/es502086e.

Thornton, D. C. O. (2014), Dissolved organic matter (DOM) release by phytoplankton in the contemporary and future ocean, *European Journal of Phycology*, *49*(1), 20-46, doi:10.1080/09670262.2013.875596.

Walker, S. A., R. M. W. Amon, C. Stedmon, S. Duan, and P. Louchouarn (2009), The use of PARAFAC modeling to trace terrestrial dissolved organic matter and fingerprint water masses in coastal Canadian Arctic surface waters, *Journal of Geophysical Research: Biogeosciences*, *114*(G4), 12 p., G00F06, doi:10.1029/2009jg000990.

Weishaar, J. L., G. R. Aiken, B. A. Bergamaschi, M. S. Fram, R. Fujii, and K. Mopper (2003), Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon, *Environmental Science & Technology*, *37*(20), 4702-4708, doi:10.1021/es030360x.

Williams, P. (1990), The importance of losses during microbial growth: Commentary on the physiology, measurement and ecology of the release of dissolved organic material, *Mar. Microb. Food Webs*, *4*(2), 175-206, Corpus ID: 89079430.

Wünsch, U. J., K. R. Murphy, and C. A. Stedmon (2017), The one-sample PARAFAC approach reveals molecular size distributions of fluorescent components in dissolved organic matter, *Environmental Science & Technology*, *51*(20), 11900-11908, doi:10.1021/acs.est.7b03260.

Xu, H., and L. Guo (2018), Intriguing changes in molecular size and composition of dissolved organic matter induced by microbial degradation and self-assembly, *Water Research*, *135*, 187-194, doi:https://doi.org/10.1016/j.watres.2018.02.016.

Yamashita, Y., F. Hashihama, H. Saito, H. Fukuda, and H. Ogawa (2017), Factors controlling the geographical distribution of fluorescent dissolved organic matter in the surface waters of the Pacific Ocean, *Limnology and Oceanography*, *62*(6), 2360-2374, doi:https://doi.org/10.1002/lno.10570.

Yamashita, Y., and E. Tanoue (2003), Chemical characterization of protein-like fluorophores in DOM in relation to aromatic amino acids, *Marine Chemistry*, *82*(3), 255-271, doi:https://doi.org/10.1016/S0304-4203(03)00073-2.

Yamashita, Y., and E. Tanoue (2004), Chemical characteristics of amino acid-containing dissolved organic matter in seawater, *Organic Geochemistry*, *35*(6), 679-692, doi:https://doi.org/10.1016/j.orggeochem.2004.02.007.

Zavorka, L., A. Crespel, N. J. Dawson, M. Papatheodoulou, S. S. Killen, and M. J. Kainz (2021), Climate change-induced deprivation of dietary essential fatty acids can reduce growth and mitochondrial efficiency of wild juvenile salmon, *Funct. Ecol.*, *35*(9), 1960-1971, doi:10.1111/1365-2435.13860.

Chapter 4: Conclusion

This thesis proposed two experiments to assess how the DOM chemical fingerprint is affected by environmental factors. In chapter 2, a chemometric method was used to delineate the mass spectra of freshwater samples, and distinguish chromophoric and fluorescent carbon. LDOM was shown to be heavier and highly selective, and its seasonal abundance was positively linked to high flow periods, suggesting an allochthonous origin. At the same time, autochthonous production in chapter 3 was enriched in carbohydrates and many kinds of FDOM moieties.

More than just a continuation, those studies completed each other with properties found in both: while C4 absorbance peaked at 340 nm and proved to be biologically labile (chapter 2), the CDOM concentration from chapter 3, estimated at 350 nm, was consumed in the early days of growth. Similarly, the C4 corresponding fluorescence at 490 nm was just as biologically degraded as the C_{510} component under the effect of microbial reprocessing. Finally, when we hypothesized in chapter 2 that a biological production of fluorescent components was able to compensate for the theoretical photolytic losses in summer, we indeed observed the production of fluorochrome in the peak A and C regions under the effect of phototroph growth (chapter 3).

Combined, my work shows a clear relation between molecular mass and biological lability: with C4 vs C2 in chapter 2, and the labile properties of diatoms' exudates associated with a lower Sr and high C_{335} for low C_{310} in chapter 3. Both findings reassert *Amon and Benner's* [1996] size-continuum theory. Additionally, while chapter 2 confirms the RDOM enrichment from the microbial loop – C2/C4 ratio – as predicted by *Jiao et al.* [2010], the activity of plankton in chapter 3 proved to preserve LDOM over a short three-week time period. Bacteria are the key converter of freshly produced DOM via heterotrophic reprocessing observed in both chapters 2 and 3.

Molecular diversity within the DOM pool spans over a continuum of molecular functions and is thought to lower its biological attractivity for reprocessing [*LaRowe et al.*, 2012]. The

branched carbon found in CRAM is both complex and hydrophobic and the impact of those physical properties remains indistinct to define a compound's lability [*Flerus et al.*, 2012]. Our results show that biological activity accommodates well a huge set of molecular weight and proton environments for degradation and production, such as complex structures of CDOM and C_{510} FDOM were quickly processed. Hence, complexity appears not to be the critical factor here. In contrast, the refractory component C2 in – chapter 2 – showed evidence of hydrophobic interaction with the column, and the C-dominated alkyl compounds targeted in chapter 3 were not very reactive. Hence, maybe the chemical surface properties of DOM aggregates have played a greater role in reactivity than chemical complexity.

In regards to biogeochemistry, Chapter 3 stressed that the current models neglect the change in DOM properties under climate change. As we showed, the proportion of heterotrophic activity and primary producers control the DOM chemical properties. Considering that the storage of carbon in the biomass is optimal when carbon reprocessing is efficient (less carbon loss during cell activity), the shift in DOM properties due to climate change will likely increase carbon respiration in the coastal areas of the North Atlantic [*Roy et al.*, 2011; *Sanders et al.*, 2014]. It seems that the carbon sequestration by the microbial carbon pump of the North Atlantic could slow down due to this effect combined with the forecasted decrease in POC fluxes [*Barange et al.*, 2017].

A fascinating highlight of Chapter 2 was that the patterns behind the immense chemical diversity of DOM can be well modeled by a small amount of hydrology and seasonal factors. Just as found by *Merder et al.* [2021], the abundance of those thousands of chemicals is not independent but covaries proportionally within the pool. While one would have assumed that parametrizing thousands of chemical behaviors in their model would be an impossible task before this work, our results showed that consistent clusters can be coined as an output for further forecasting. It seems that only a handful of new terms should already produce a satisfying estimate if a well-designed

experiment was made to quantify their association with changes in BGE, energetic density, POC formation potential, and respiration rates. Effectively, chemometrics approaches, like PARAFAC, are designed to co-cluster non-independent behaviors and therefore can digest those massive datasets into systems humanly understandable. A bit like sociology has progressed by privileging trends upon individuals' opinions, maybe a fingerprinting approach to DOM in which analytes are categorized into groups of behaviors could overcome the arising limitation of data complexity.

Monitoring only DOM fluxes while considering it holds a stable identity seems every day a less adequate assumption to make [*Anderson et al.*, 2015]. Evidence from both experiments, and others, demonstrate that the relative abundance of CDOM, FDOM, RDOM, and LDOM fluctuates through seasons; that the autochthonous production diverges given ecological niche; and that those fresh nutrients can induce the degradation of material initially refractory. DOM holds a high energy density but is limited by hetero-atoms scarcity. An evolutional perspective suggests that, given the appropriate strategy, any accumulating resource will eventually be used as the ground for an ecological niche. Hence, recalcitrance is context-dependent stressed *Shen and Benner* [2018], and the coupling of inorganic nutrients to this material appears as the 'adequate strategy' for microbial utilization, therefore debunking the concept of intrinsic stability.

Reference :

Abdulla, H., E. C. Minor, R. F. Dias, and P. G. Hatcher (2010), Changes in the compound classes of dissolved organic matter along an estuarine transect: A study using FTIR and 13C NMR, *Geochimica et Cosmochimica Acta*, *74*(13), 3815-3838, doi:https://doi.org/10.1016/j.gca.2010.04.006.

Amaral, V., D. Graeber, D. Calliari, and C. Alonso (2016), Strong linkages between DOM optical properties and main clades of aquatic bacteria, *Limnology and Oceanography*, *61*(3), 906-918, doi:10.1002/lno.10258.

Amon, R., and R. Benner (1996), Bacterial utilization of different size classes of dissolved organic matter, *Limnology and Oceanography*, *41*(1), 41-51, doi:https://doi.org/10.4319/lo.1996.41.1.0041.

Anderson, L. G., and R. M. W. Amon (2015), Chapter 14 - DOM in the Arctic Ocean, in *Biogeochemistry of marine dissolved organic matter (second edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 609-633, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00014-5.

Anderson, T. R., J. R. Christian, and K. J. Flynn (2015), Chapter 15 - Modeling DOM biogeochemistry, in *Biogeochemistry of marine dissolved organic matter (second edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 635-667, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00015-7.

Ateia, M., J. Ran, M. Fujii, and C. Yoshimura (2017), The relationship between molecular composition and fluorescence properties of humic substances, *International Journal of Environmental Science and Technology*, *14*(4), 867-880, doi:10.1007/s13762-016-1214-x.

Balaguru, K., S. C. Doney, L. Bianucci, P. J. Rasch, L. R. Leung, J. H. Yoon, and I. D. Lima (2018), Linking deep convection and phytoplankton blooms in the northern Labrador Sea in a changing climate, *PLoS One*, *13*(1), 17 p., doi:10.1371/journal.pone.0191509.

Barange, M., M. Butenschon, A. Yool, N. Beaumont, J. A. Fernandes, A. P. Martin, and J. I. Allen (2017), The cost of reducing the north Atlantic Ocean biological carbon pump, *Frontiers in Marine Science*, *3*, 10, doi:10.3389/fmars.2016.00290.

Beaupré, S. R. (2015), Chapter 6 - The carbon isotopic composition of marine DOC, in *Biogeochemistry of marine dissolved organic matter (second edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 335-368, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00006-6.

Benavides, M., and M. Voss (2015), Five decades of N2 fixation research in the North Atlantic Ocean, *Frontiers in Marine Science*, *2*(40), 20 p., doi:10.3389/fmars.2015.00040.

Berman-Frank, I., and Z. Dubinsky (1999), Phytoplankton use the imbalance between carbon assimilation and biomass production to their strategic advantage balanced Growth in Aquatic Plants: Myth or Reality?, *BioScience*, *49*(1), 29-37, doi:10.1525/bisi.1999.49.1.29.

Bjørrisen, P. K. (1988), Phytoplankton exudation of organic matter: Why do healthy cells do it?, *Limnology and Oceanography*, *33*(1), 151-154, doi:https://doi.org/10.4319/lo.1988.33.1.0151.

Brown, S. B., J. A. Holroyd, D. I. Vernon, Y. K. Shim, and K. M. Smith (1989), The biosynthesis of the chromophore of phycocyanin. Pathway of reduction of biliverdin to phycocyanobilin, *The Biochemical journal*, *261*(1), 259-263, doi:10.1042/bj2610259.

Buck, K. R., F. P. Chavez, and L. Campbell (1996), Basin-wide distributions of living carbon components and the inverted trophic pyramid of the central gyre of the North Atlantic Ocean, summer 1993, *Aquatic Microbial Ecology*, *10*(3), 283-298, doi:10.3354/ame010283.

Burkhard, L. P. (2000), Estimating dissolved organic carbon partition coefficients for non-ionic organic chemicals, *Environmental Science & Technology*, *34*(22), 4663-4668, doi:10.1021/es001269l.

Carlson, C. A., and D. A. Hansell (2015), Chapter 3 - DOM sources, sinks, reactivity, and budgets, in *Biogeochemistry of marine dissolved organic matter (second edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 65-126, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00003-0.

Catalá, T. S., S. Shorte, and T. Dittmar (2021), Marine dissolved organic matter: a vast and unexplored molecular space, *Applied Microbiology and Biotechnology*, *105*(19), 7225-7239, doi:10.1007/s00253-021-11489-3.

Cauwet, G. (1999), Determination of dissolved organic carbon and nitrogen by high temperature combustion, in *Methods of Seawater Analysis*, edited, pp. 407-420, doi:10.1002/9783527613984.ch15.

Chen, M., S. Kim, J. E. Park, H. J. Jung, and J. Hur (2016), Structural and compositional changes of dissolved organic matter upon solid-phase extraction tracked by multiple analytical tools, *Analytical and Bioanalytical Chemistry*, *408*(23), 6249-6258, doi:10.1007/s00216-016-9728-0.

Coble, P. G., J. Lead, A. Baker, D. M. Reynolds, and R. Spencer (2014), *Aquatic organic matter fluorescence*, 375 pp., Cambridge University Press, Cambridge, doi:10.1017/CBO9781139045452.

Cooper, L. W., R. Benner, J. W. McClelland, B. J. Peterson, R. M. Holmes, P. A. Raymond, D. A. Hansell, J. M. Grebmeier, and L. A. Codispoti (2005), Linkages among runoff, dissolved organic carbon, and the stable oxygen isotope composition of seawater and other water mass indicators in the Arctic Ocean, *Journal of Geophysical Research: Biogeosciences*, *110*(G2), 14 p., doi:https://doi.org/10.1029/2005JG000031.

Cooper, W., J. Chanton, J. D'Andrilli, S. Hodgkins, D. Podgorski, A. Stenson, M. Tfaily, and R. Wilson (2020), A History of molecular level analysis of natural organic matter by FTICR mass spectrometry and the paradigm shift in organic geochemistry, *Mass Spectrometry Reviews*, *41*(2), 215-239, doi:10.1002/mas.21663.

Coutinho, F., P. Meirelles, A. P. Moreira, R. Paranhos, B. Dutilh, and F. Thompson (2015), Niche distribution and influence of environmental parameters in marine microbial communities: a systematic review, *PeerJ*, *3*, 19 p., doi:10.7717/peerj.1008.

Daniels, C., A. Poulton, M. Esposito, M. Paulsen, R. Bellerby, A. Martin, and M. John (2015), Phytoplankton dynamics in contrasting early stage North Atlantic spring blooms: Composition, succession, and potential drivers, *Biogeosciences*, *12*(8), 2395-2409, doi:10.5194/bg-12-2395-2015.

De La Rocha, C. L., and U. Passow (2014), Section 4 - The biological pump, in *Treatise on Geochemistry*, edited by H. D. Holland and K. K. Turekian, pp. 93-112, Volume 8, 2nd edition, Oxford: Elsevier, doi:http://dx.doi.org/10.1016/B978-0-08-095975-7.00604-5.

Dell'Anno, A., and R. Danovaro (2005), Extracellular DNA plays a key role in deep-Sea ecosystem functioning, *Science*, *309*(5744), 2179-2179, doi:10.1126/science.1117475.

Derrien, M., S. R. Brogi, and R. Goncalves-Araujo (2019), Characterization of aquatic organic matter: Assessment, perspectives and research priorities, *Water Res*, *163, 114908*, 17 p., doi:10.1016/j.watres.2019.114908.

Ducklow, H. W. (1999), The bacterial component of the oceanic euphotic zone, *Microbiology Ecology*, *30*(1), 1-10, doi:https://doi.org/10.1111/j.1574-6941.1999.tb00630.x.

Duursma, E. K. (1963), The production of dissolved organic matter in the sea, as related to the primary gross production of organic matter, *Netherlands Journal of Sea Research*, *2*(1), 85-94, doi:https://doi.org/10.1016/0077-7579(63)90007-3.

Estrada, E., E. J. Delgado, and Y. Simón-Manso (2007), Chapter 2 - Modeling the solubility in water of environmentally important organic compounds, in *Thermodynamics, Solubility and Environmental Issues*, edited by T. M. Letcher, pp. 17-31, Elsevier, Amsterdam, doi:https://doi.org/10.1016/B978-044452707-3/50004-5.

Farag, M. A., A. Meyer, S. E. Ali, M. A. Salem, P. Giavalisco, H. Westphal, and L. A. Wessjohann (2018), Comparative metabolomics approach detects stress-specific responses during coral bleaching in soft corals, *J Proteome Res*, *17*(6), 2060-2071, doi:10.1021/acs.jproteome.7b00929.

Fitznar, H. P., J. M. Lobbes, and G. Kattner (1999), Determination of enantiomeric amino acids with high-performance liquid chromatography and pre-column derivatisation with o-phthaldialdehyde and N-isobutyrylcysteine in seawater and fossil samples (mollusks), *Journal of Chromatography A*, *832*(1), 123-132, doi:https://doi.org/10.1016/S0021-9673(98)01000-0.

Flerus, R., O. Lechtenfeld, B. Koch, S. McCallister, P. Schmitt-Kopplin, R. Benner, K. Kaiser, and G. Kattner (2012), A molecular perspective on the ageing of marine dissolved organic matter, *Biogeosciences*, *9*(6), 1935–1955, doi:10.5194/bg-9-1935-2012.

Fogg, G. E. (1966), The extracellular products of algae, *Oceanogr. Mar. Biol. Ann. Rev.*, *4*, 195-212.

Gibson, J., and C. S. Harwood (2002), Metabolic diversity in aromatic compound utilization by anaerobic microbes, *Annual Review of Microbiology*, *56*(1), 345-369, doi:10.1146/annurev.micro.56.012302.160749.

Giorgio, P. A. d., and J. J. Cole (1998), Bacterial growth efficiency in natural aquatic systems, *Annual Review of Ecology and Systematics*, *29*(1), 503-541, doi:10.1146/annurev.ecolsys.29.1.503.

Greene, C. H., A. J. Pershing, T. M. Cronin, and N. Ceci (2008), Arctic climatechange and its impacts on the ecology of the North Atlantic, *Ecology*, *89*(11 S), 24-38, doi:10.1890/07-0550.1.

Gueguen, C., and P. Kowalczuk (2014), Colored dissolved organic matter in frontal zones, in *The handbook of environmental chemistry* edited by D. Barceló and A. G. Kostianoy, pp. 1-35, Springer, Berlin, doi:10.1007/698_2013_244.

Guruprasad, S., and T. V. Ramachandra (2011), Chronicle of marine diatom culturing techniques, *Indian Journal of Fundamental and Applied Life Sciences*, *1*(3), 282-294. URL: http://eprints.iisc.ac.in/id/eprint/42805.

Hansell, D. A., and C. A. Carlson (2014), *Biogeochemistry of marine dissolved organic matter*, 698 pp., Elsevier Science & Technology, San Diego, US, ISBN 9780124071537.

Hawkes, J. A., et al. (2015), Efficient removal of recalcitrant deep-ocean dissolved organic matter during hydrothermal circulation, *Nature Geoscience*, *8*(11), 856-860, doi:10.1038/ngeo2543.

Hawkes, J. A., P. J. Sjöberg, J. Bergquist, and L. J. Tranvik (2019), Complexity of dissolved organic matter in the molecular size dimension: insights from coupled size exclusion chromatography electrospray ionisation mass spectrometry, *Faraday Discussions*, *218*, 52-71, doi:10.1039/C8FD00222C.

Heal, K. R., B. Durham, A. K. Boysen, L. T. Carlson, W. Qin, F. Ribalet, A. E. White, R. M. Bundy, E. V. Armbrust, and A. E. Ingalls (2020), Marine community metabolomes carry fingerprints of phytoplankton community composition, *bioRxiv*, *6*(3), 19 p., doi:10.1101/2020.12.22.424086.

Hébert, M., and L. Tremblay (2017), Production and persistence of bacterial and labile organic matter at the hypoxic water–sediment interface of the St. Lawrence Estuary, *Limnology and Oceanography*, *62*(5), 2154-2167, doi:https://doi.org/10.1002/lno.10556.

Helms, J. R., J. Mao, K. Schmidt-Rohr, H. Abdulla, and K. Mopper (2013), Photochemical flocculation of terrestrial dissolved organic matter and iron, *Geochimica et Cosmochimica Acta*, *121*, 398-413, doi:https://doi.org/10.1016/j.gca.2013.07.025.

Helms, J. R., A. Stubbins, J. D. Ritchie, E. C. Minor, D. J. Kieber, and K. Mopper (2008), Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter, *Limnology and Oceanography*, *53*(3), 955-969, doi:10.4319/lo.2008.53.3.0955.

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*(3), 1583-1624, doi:10.5194/bg-10-1583-2013.

Hirose, K. (2007), Metal–organic matter interaction: Ecological roles of ligands in oceanic DOM, *Applied Geochemistry*, *22*(8), 1636-1645, doi:https://doi.org/10.1016/j.apgeochem.2007.03.042.

Huuskonen, J. (2000), Estimation of aqueous solubility for a diverse set of organic compounds based on molecular topology, *Journal of Chemical Information and Computer Sciences*, *40*(3), 773-777, doi:10.1021/ci9901338.

Jessen, G. L., A. Lichtschlag, A. Ramette, S. Pantoja, P. E. Rossel, C. J. Schubert, U. Struck, and A. Boetius (2017), Hypoxia causes preservation of labile organic matter and changes seafloor microbial community composition (Black Sea), *Science Advances*, *3*(2), 15 p., doi:10.1126/sciadv.1601897.

Jiao, N., et al. (2010), Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean, *Nature Reviews Microbiology*, *8*(8), 593-599, doi:10.1038/nrmicro2386.

Jørgensen, L., C. A. Stedmon, M. A. Granskog, and M. Middelboe (2014), Tracing the long-term microbial production of recalcitrant fluorescent dissolved organic matter in seawater, *Geophysical Research Letters*, *41*(7), 2481-2488, doi:https://doi.org/10.1002/2014GL059428.

Keller, D. P., and R. R. Hood (2011), Modeling the seasonal autochthonous sources of dissolved organic carbon and nitrogen in the upper Chesapeake Bay, *Ecological Modelling*, *222*(5), 1139-1162, doi:https://doi.org/10.1016/j.ecolmodel.2010.12.014.

Kennicutt, M. C., and L. M. Jeffrey (1981), Chemical and GC-MS characterization of marine dissolved lipids, *Marine Chemistry*, *10*(5), 367-387, doi:https://doi.org/10.1016/0304-4203(81)90016-5.

Kieber, D. J., J. McDaniel, and K. Mopper (1989), Photochemical source of biological substrates in sea water: implications for carbon cycling, *Nature*, *341*(6243), 637-639, doi:10.1038/341637a0.

Kothawala, D. N., C. A. Stedmon, R. A. Müller, G. A. Weyhenmeyer, S. J. Köhler, and L. J. Tranvik (2014), Controls of dissolved organic matter quality: evidence from a large-scale boreal lake survey, *Global Change Biology*, *20*(4), 1101-1114, doi:https://doi.org/10.1111/gcb.12488.

Kowalczuk, P., G. H. Tilstone, M. Zabłocka, R. Röttgers, and R. Thomas (2013), Composition of dissolved organic matter along an Atlantic Meridional Transect from fluorescence spectroscopy and Parallel Factor Analysis, *Marine Chemistry*, *157*, 170-184, doi:https://doi.org/10.1016/j.marchem.2013.10.004.

Kujawinski, E. B., R. Del Vecchio, N. V. Blough, G. C. Klein, and A. G. Marshall (2004), Probing molecular-level transformations of dissolved organic matter: insights on photochemical

degradation and protozoan modification of DOM from electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, *Marine Chemistry*, *92*(1), 23-37, doi:https://doi.org/10.1016/j.marchem.2004.06.038.

Kumar, K., R. A. Mella-Herrera, and J. W. Golden (2010), Cyanobacterial heterocysts, *Cold Spring Harb Perspect Biol*, *2*(4), 19 p., doi:10.1101/cshperspect.a000315.

Lam, B., and A. J. Simpson (2008), Direct 1H NMR spectroscopy of dissolved organic matter in natural waters, *Analyst*, *133*(2), 263-269, doi:10.1039/B713457F.

LaRowe, D. E., A. W. Dale, J. P. Amend, and P. Van Cappellen (2012), Thermodynamic limitations on microbially catalyzed reaction rates, *Geochimica et Cosmochimica Acta*, *90*, 96-109, doi:https://doi.org/10.1016/j.gca.2012.05.011.

Lechtenfeld, 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, doi:https://doi.org/10.1016/j.gca.2013.11.009.

Lechtenfeld, O. J., N. Hertkorn, Y. Shen, M. Witt, and R. Benner (2015), Marine sequestration of carbon in bacterial metabolites, *Nat Commun*, *6*(1), 6711, doi:10.1038/ncomms7711.

Lehmann, J., and M. Kleber (2015), The contentious nature of soil organic matter, *Nature*, *528*(7580), 60-68, doi:10.1038/nature16069.

Lennartz, S. T., and T. Dittmar (2022), Controls on turnover of marine dissolved organic matter—testing the null hypothesis of purely concentration-driven uptake: Comment on Shen and Benner, "Molecular properties are a primary control on the microbial utilization of dissolved organic matter in the ocean", *Limnology and Oceanography*, *67*(3), 673-679, doi:https://doi.org/10.1002/lno.12028.

Leyva, D., M. U. Tariq, R. Jaffé, F. Saeed, and F. F. Lima (2022), Unsupervised structural classification of dissolved organic matter based on fragmentation pathways, *Environmental Science & Technology*, *56*(2), 1458-1468, doi:10.1021/acs.est.1c04726.

Lynch, L. M., N. A. Sutfin, T. S. Fegel, C. M. Boot, T. P. Covino, and M. D. Wallenstein (2019), River channel connectivity shifts metabolite composition and dissolved organic matter chemistry, *Nature Communications*, *10*(1), 459, doi:10.1038/s41467-019-08406-8.

Mangal, V., N. L. Stock, and C. Guéguen (2016), Molecular characterization of phytoplankton dissolved organic matter (DOM) and sulfur components using high resolution Orbitrap mass spectrometry, *Analytical and Bioanalytical Chemistry*, *408*(7), 1891-1900, doi:10.1007/s00216-015-9295-9.

Mannino, A., et al. (2019), Volume 5.0- Measurement protocol of absorption by chromophoric dissolved organic matter (CDOM) and other dissolved materials, in *Ocean optics and biogeochemistry protocols for satellite ocean colour sensor validation*, edited by A. Mannino and

M. G. Novak, International ocean coordinating group, Dartmouth, Canada, doi:http://dx.doi.org/10.25607/OBP-119.

Merder, J., H. Röder, T. Dittmar, U. Feudel, J. Freund, G. Gerdts, A. Kraberg, and J. Niggemann (2021), Dissolved organic compounds with synchronous dynamics share chemical properties and origin, *Limnology and Oceanography*, *66*(11), 4001-4016, doi:10.1002/lno.11938.

Messetta, M. L., C. Hegoburu, J. P. Casas-Ruiz, A. Butturini, and C. Feijoó (2018), Characterization and qualitative changes in DOM chemical characteristics related to hydrologic conditions in a Pampean stream, *Hydrobiologia*, *808*(1), 201-217, doi:10.1007/s10750-017-3422-x.

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*(6), 1317-1324, doi:https://doi.org/10.4319/lo.1997.42.6.1317.

Mitra, A., et al. (2014), The role of mixotrophic protists in the biological carbon pump, *Biogeosciences*, *11*(4), 995-1005, doi:10.5194/bg-11-995-2014.

Mopper, K., D. J. Kieber, and A. Stubbins (2015), Chapter 8 - Marine photochemistry of organic matter: processes and impacts, in *Biogeochemistry of marine dissolved organic matter (Second Edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 389-450, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00008-X.

Moran, M. A., and B. P. Durham (2019), Sulfur metabolites in the pelagic ocean, *Nature Reviews Microbiology*, *17*(11), 665-678, doi:10.1038/s41579-019-0250-1.

Murphy, K. R., C. A. Stedmon, D. Graeber, and R. Bro (2013), Fluorescence spectroscopy and multi-way techniques: PARAFAC, *Analytical Methods*, *5*(23), 6557-6566, doi:10.1039/C3AY41160E.

Murphy, K. R., C. A. Stedmon, T. D. Waite, and G. M. Ruiz (2008), Distinguishing between terrestrial and autochthonous organic matter sources in marine environments using fluorescence spectroscopy, *Marine Chemistry*, *108*(1), 40-58, doi:https://doi.org/10.1016/j.marchem.2007.10.003.

Murphy, K. R., C. A. Stedmon, P. Wenig, and R. Bro (2014), OpenFluor- an online spectral library of auto-fluorescence by organic compounds in the environment, *Analytical Methods*, *6*(3), 658-661, doi:10.1039/C3AY41935E.

Nebbioso, A., P. Mazzei, and D. Savy (2014), Reduced complexity of multidimensional and diffusion NMR spectra of soil humic fractions as simplified by Humeomics, *Chemical and Biological Technologies in Agriculture*, *1*(1), 24, doi:10.1186/s40538-014-0024-y.

Nieto-Cid, M., X. A. Álvarez-Salgado, J. Gago, and F. F. Pérez (2005), DOM fluorescence, a tracer for biogeochemical processes in a coastal upwelling system (NW Iberian Peninsula), *Marine Ecology Progress Series*, *297*, 33-50, URL: http://www.jstor.org/stable/24868645.

Noriega-Ortega, B. E., G. Wienhausen, A. Mentges, T. Dittmar, M. Simon, and J. Niggemann (2019), Does the chemodiversity of bacterial exometabolomes sustain the chemodiversity of marine dissolved organic matter?, *Frontiers in Microbiology*, *10*(215), 13 p., doi:https://doi.org/10.3389/fmicb.2019.00215.

Ogawa, H., and E. Tanoue (2003), Dissolved organic matter in oceanic waters, *Journal of Oceanography*, *59*(2), 129-147, doi:10.1023/A:1025528919771.

Pershing, A. J., and K. Stamieszkin (2019), The north Atlantic ecosystem, from plankton to whales, *Ann Rev Mar Sci*, *12*(1), 339-359, doi:10.1146/annurev-marine-010419-010752.

Pisani, O., S. D. Frey, A. J. Simpson, and M. J. Simpson (2015), Soil warming and nitrogen deposition alter soil organic matter composition at the molecular-level, *Biogeochemistry*, *123*(3), 391-409, doi:10.1007/s10533-015-0073-8.

Raymond, P. A., and R. G. M. Spencer (2015), Chapter 11 - Riverine DOM, in *Biogeochemistry of marine dissolved organic matter (second edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 509-533, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00011-X.

Reader, H. E., and W. L. Miller (2012), Variability of carbon monoxide and carbon dioxide apparent quantum yield spectra in three coastal estuaries of the South Atlantic Bight, *Biogeosciences*, *9*(11), 4279-4294, doi:10.5194/bg-9-4279-2012.

Reader, H. E., C. A. Stedmon, and E. S. Kritzberg (2014), Seasonal contribution of terrestrial organic matter and biological oxygen demand to the Baltic Sea from three contrasting river catchments, *Biogeosciences*, *11*(12), 3409-3419, doi:10.5194/bg-11-3409-2014.

Reader, H. E., C. A. Stedmon, N. J. Nielsen, and E. S. Kritzberg (2015), Mass and UV-visible spectral fingerprints of dissolved organic matter: sources and reactivity, *Frontiers in Marine Science*, *2*(88), 10 p., doi:10.3389/fmars.2015.00088.

Redfield, A. C. (1958), The biological control of chemical factors in the environment, *Science progress*, *11*, 150-170, PMID: 24545739.

Reinthaler, T., X. A. Á. Salgado, M. Alvarez, H. M. van Aken, and G. J. Herndl (2013), Impact of water mass mixing on the biogeochemistry and microbiology of the Northeast Atlantic Deep Water, *Global biogeochemical cycles*, *27*(4), 1151-1162, doi:10.1002/2013gb004634.

Repeta, D. J. (2015), Chapter 2 - Chemical characterization and cycling of dissolved organic matter, in *Biogeochemistry of marine dissolved organic matter (second edition)*, edited by D. A. Hansell and C. A. Carlson, pp. 21-63, Academic Press, Boston, doi:https://doi.org/10.1016/B978-0-12-405940-5.00002-9.

Richardson, T., and J. Cullen (1995), Changes in buoyancy and chemical composition during growth of a coastal marine diatom: Ecological and biogeochemical consequences, *Marine Ecology Progress Series*, *128*(1), 77-90, doi:10.3354/meps128077.

Rolland, J. L., D. Stien, S. Sanchez-Ferandin, and R. Lami (2016), Quorum sensing and quorum quenching in the phycosphere of phytoplankton: a case of chemical interactions in ecology, *J Chem Ecol*, *42*(12), 1201-1211, doi:10.1007/s10886-016-0791-y.

Roy, T., L. Bopp, M. Gehlen, B. Schneider, P. Cadule, T. L. Frolicher, J. Segschneider, J. Tjiputra, C. Heinze, and F. Joos (2011), Regional impacts of climate change and atmospheric CO2 on future ocean carbon uptake: a multimodel linear feedback analysis, *J. Clim.*, *24*(9), 2300-2318, doi:10.1175/2010jcli3787.1.

Ruess, L., and D. C. Muller-Navarra (2019), Essential biomolecules in food webs, *Front. Ecol. Evol.*, *7*(269), 18, doi:10.3389/fevo.2019.00269.

Sanders, R., et al. (2014), The biological carbon pump in the north Atlantic, *Progress in Oceanography*, *129*(part B), 200-218, doi:https://doi.org/10.1016/j.pocean.2014.05.005.

Sandron, S., A. Rojas, R. Wilson, N. W. Davies, P. R. Haddad, R. A. Shellie, P. N. Nesterenko, B. P. Kelleher, and B. Paull (2015), Chromatographic methods for the isolation, separation and characterisation of dissolved organic matter, *Environmental Science: Processes & Impacts*, *17*(9), 1531-1567, doi:10.1039/C5EM00223K.

Seth, K., G. Kumawat, M. Kumar, V. Sangela, N. Singh, A. K. Gupta, and Harish (2021), Chapt. 12- Nitrogen metabolism in cyanobacteria, in *Ecophysiology and biochemistry of cyanobacteria*, edited by R. P. Rastogi, pp. 255-268, 1st ed., Springer, Singapore, doi:10.1007/978-981-16-4873-1_12.

She, Z., J. Wang, C. He, X. Pan, Y. Li, S. Zhang, Q. Shi, and Z. Yue (2021), The stratified distribution of dissolved organic matter in an AMD lake revealed by multi-sample evaluation procedure, *Environ Sci Technol*, *55*(12), 8401-8409, doi:10.1021/acs.est.0c05319.

Shen, Y., and R. Benner (2018), Mixing it up in the ocean carbon cycle and the removal of refractory dissolved organic carbon, *Scientific Reports*, *8*(1), 2542, doi:10.1038/s41598-018-20857-5.

Smith, D. C., M. Simon, A. L. Alldredge, and F. Azam (1992), Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution, *Nature*, *359*(6391), 139-142, doi:10.1038/359139a0.

Solomon, C. T., et al. (2015), Ecosystem consequences of changing inputs of terrestrial dissolved organic matter to lakes: current knowledge and future challenges, *Ecosystems*, *18*(3), 376-389, doi:10.1007/s10021-015-9848-y.

Song, L. (2017), A multiomics approach to study the microbiome response to phytoplankton blooms, *Appl Microbiol Biotechnol*, *101*(12), 4863-4870, doi:10.1007/s00253-017-8330-5.

Stryer, L., J. M. Berg, and J. L. Tymoczko (2013), Partie II - La transduction et la mise en réserve de l'énergie -Chapt. 15, 16, 22 and 23, in *Biochimie*, edited by Medecine science publications, p. 1104, Lavoisier, 7 ed., ISBN 9782257204271.

Stubbins, A., J. F. Lapierre, M. Berggren, Y. T. Prairie, T. Dittmar, and P. A. del Giorgio (2014), What's in an EEM? Molecular signatures associated with dissolved organic fluorescence in boreal Canada, *Environ Sci Technol*, *48*(18), 10598-10606, doi:10.1021/es502086e.

Taipale, S. J., M. Hiltunen, K. Vuorio, and E. Peltomaa (2016), Suitability of phytosterols alongside fatty acids as chemotaxonomic biomarkers for phytoplankton, *Frontiers in Plant Science*, *7*(212), 16 p., doi:10.3389/fpls.2016.00212.

Valeur, B., and M. N. Berberan-Santos (2012a), Chapter 4- Structural effects on fluorescence emission, in *Molecular fluorescence: principles and applications, Second edition*, edited, pp. 75-107, Wiley-VCH Verlag GmbH & Co. KGaA, doi:10.1002/9783527650002.ch4.

Valeur, B., and M. N. Berberan-Santos (2012b), Chapter 5- Environmental effects on fluorescence emission, in *Molecular fluorescence: principles and applications, Second edition*, edited, pp. 109-140, Wiley-VCH Verlag GmbH & Co. KGaA, doi:10.1002/9783527650002.ch5.

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, *Marine Chemistry*, *92*(1), 67-85, doi:https://doi.org/10.1016/j.marchem.2004.06.017.

Verdugo, P., and P. Santschi (2010), Polymer dynamics of DOC networks and gel formation in seawater, *Deep Sea Research Part II: Topical Studies in Oceanography*, *57*(16), 1486-1493, doi:10.1016/j.dsr2.2010.03.002.

Vidon, P., L. E. Wagner, and E. Soyeux (2008), Changes in the character of DOC in streams during storms in two midwestern watersheds with contrasting land uses, *Biogeochemistry*, *88*(3), 257-270, URL: https://www.jstor.org/stable/40343578.

Vlassov, V. V., P. P. Laktionov, and E. Y. Rykova (2007), Extracellular nucleic acids, *Bioessays*, *29*(7), 654-667, doi:10.1002/bies.20604.

Wagner, S., et al. (2020), Soothsaying DOM: A current perspective on the future of oceanic dissolved organic carbon, *Frontiers in Marine Science*, *7*(341), 17, doi:10.3389/fmars.2020.00341.

Walker, S. A., R. M. W. Amon, C. Stedmon, S. Duan, and P. Louchouarn (2009), The use of PARAFAC modeling to trace terrestrial dissolved organic matter and fingerprint water masses in coastal Canadian Arctic surface waters, *Journal of Geophysical Research: Biogeosciences*, *114*(G4), 12 p., G00F06, doi:10.1029/2009jg000990.

Weishaar, J. L., G. R. Aiken, B. A. Bergamaschi, M. S. Fram, R. Fujii, and K. Mopper (2003), Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon, *Environmental Science & Technology*, *37*(20), 4702-4708, doi:10.1021/es030360x.

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*(10), 781-788, doi:10.2307/1313569.

Williams, P. (1990), The importance of losses during microbial growth: Commentary on the physiology, measurement and ecology of the release of dissolved organic material, *Mar. Microb. Food Webs*, *4*(2), 175-206, Corpus ID: 89079430.

Wünsch, U., K. Murphy, and C. Stedmon (2015), Fluorescence quantum yields of natural organic matter and organic compounds: implications for the fluorescence-based interpretation of organic matter composition, *Frontiers in Marine Science*, *2*, 1-15, doi:10.3389/fmars.2015.00098.

Wünsch, U. J., and J. A. Hawkes (2020), Mathematical chromatography deciphers the molecular fingerprints of dissolved organic matter, *Analyst*, *145*(5), 1789-1800, doi:10.1039/C9AN02176K.

Wünsch, U. J., K. R. Murphy, and C. A. Stedmon (2017), The one-sample PARAFAC approach reveals molecular size distributions of fluorescent components in dissolved organic matter, *Environmental Science & Technology*, *51*(20), 11900-11908, doi:10.1021/acs.est.7b03260.

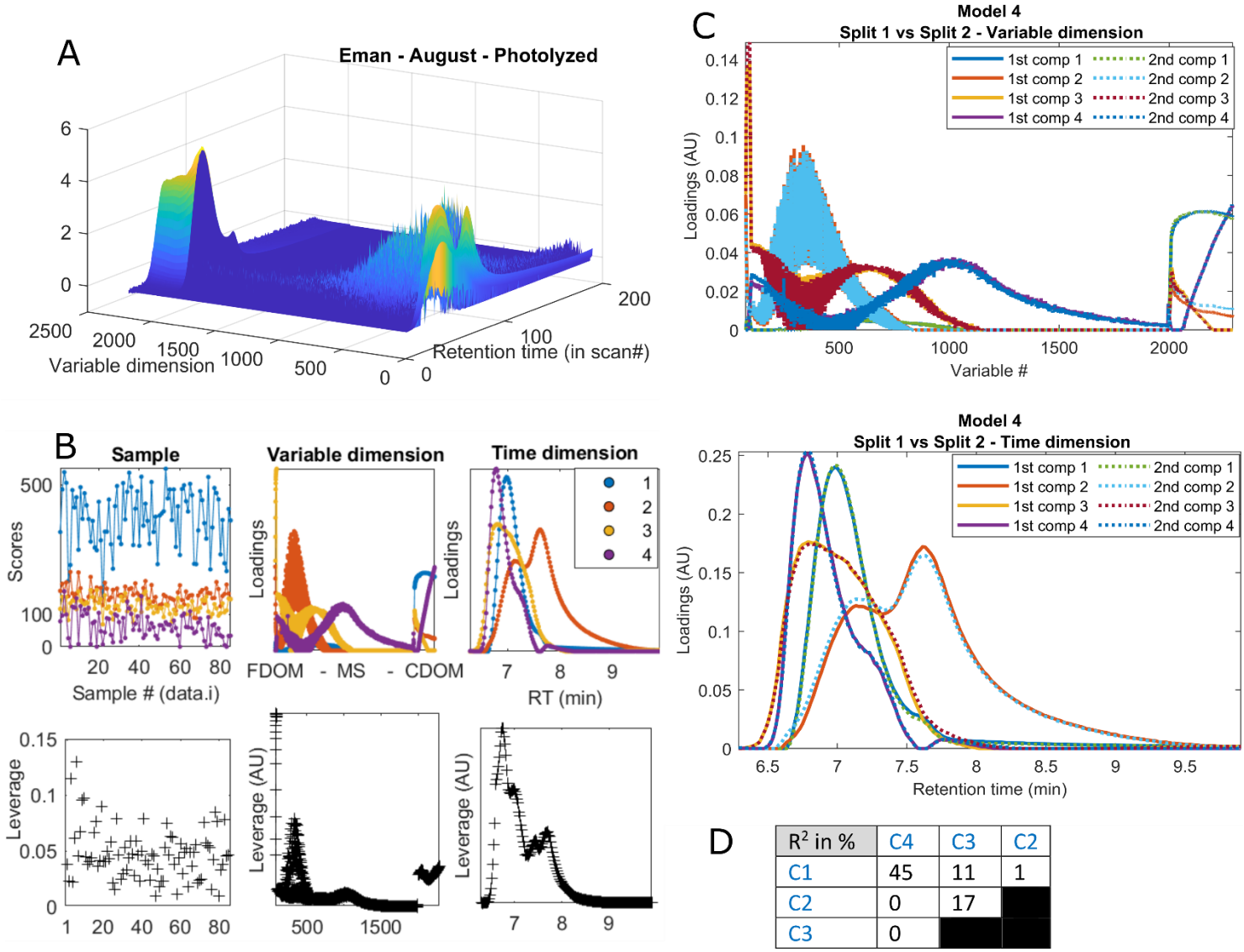
Yamashita, Y., F. Hashihama, H. Saito, H. Fukuda, and H. Ogawa (2017), Factors controlling the geographical distribution of fluorescent dissolved organic matter in the surface waters of the Pacific Ocean, *Limnology and Oceanography*, *62*(6), 2360-2374, doi:https://doi.org/10.1002/lno.10570.

Zhang, Q., R. M. Warwick, C. L. McNeill, C. E. Widdicombe, A. Sheehan, and S. Widdicombe (2015), An unusually large phytoplankton spring bloom drives rapid changes in benthic diversity and ecosystem function, *Progress in Oceanography*, *137*(part B), 533-545, doi:https://doi.org/10.1016/j.pocean.2015.04.029.

Zhang, Y., Y. Su, Z. Liu, J. Yu, and M. Jin (2017), Sediment lipid biomarkers record phytoplankton dynamics of Lake Heihai (Yunnan Province, SW China) driven by climate warming since the 1980s, *Environmental Science and Pollution Research*, *24*(26), 21509-21516, doi:10.1007/s11356-017-9931-3.

Zhuang, W.-E., and L. Yang (2018), Impacts of global changes on the biogeochemistry and environmental effects of dissolved organic matter at the land-ocean interface: a review, *Environmental Science and Pollution Research*, *25*(5), 4165-4173, doi:10.1007/s11356-017-1027-6.

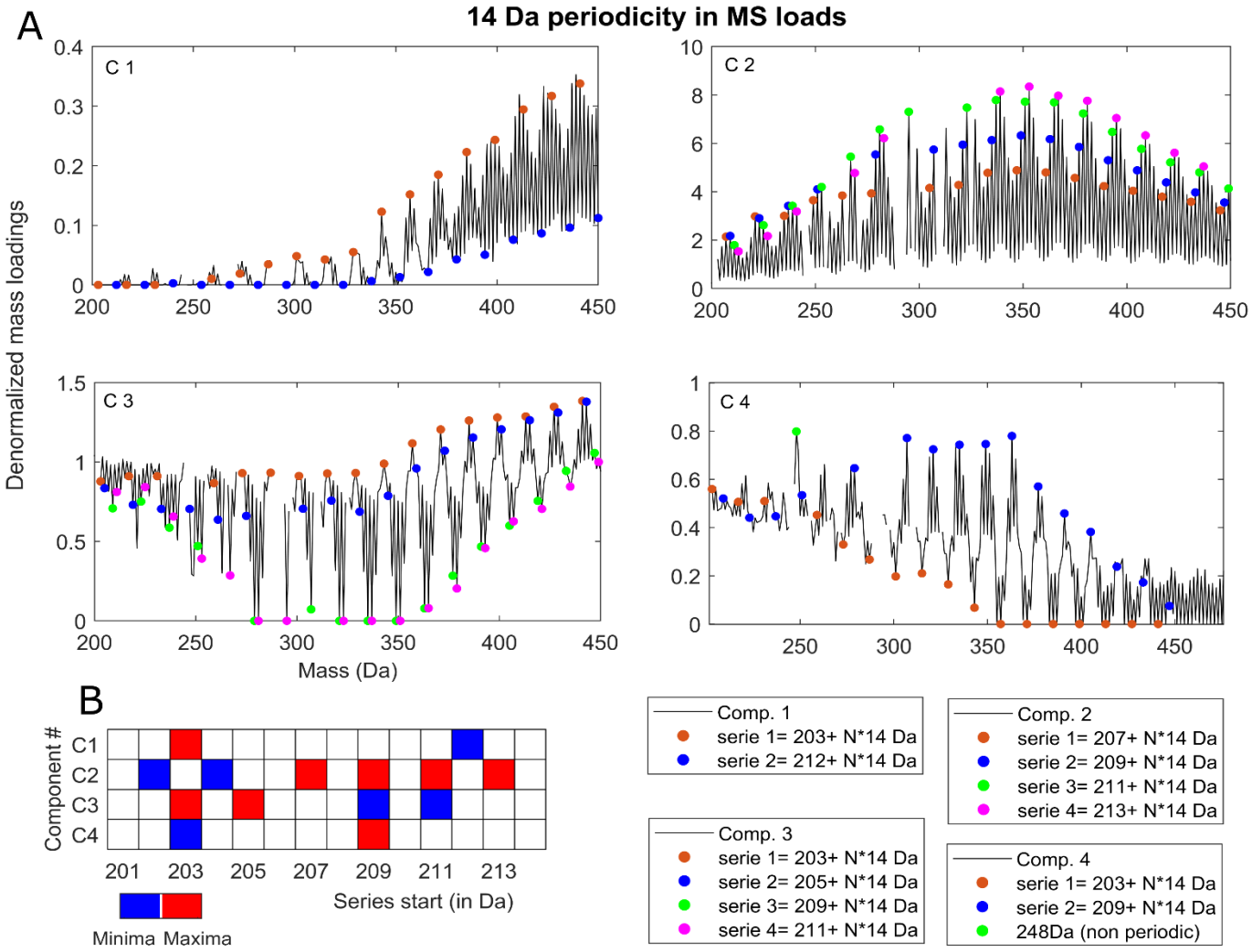
Appendices:



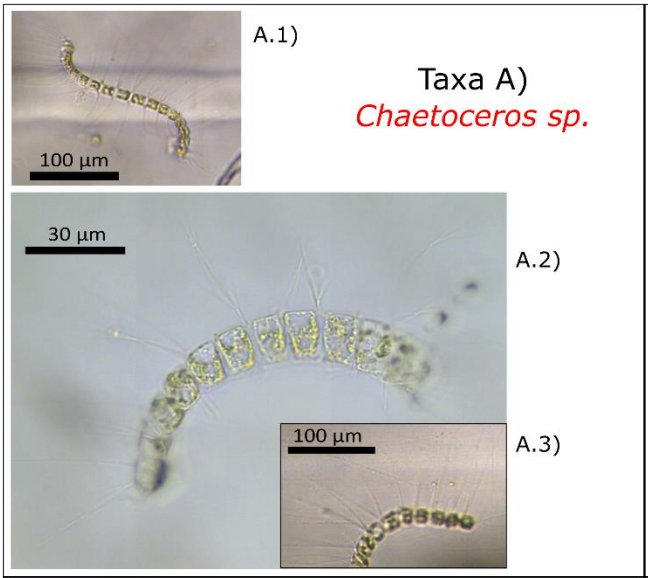
D

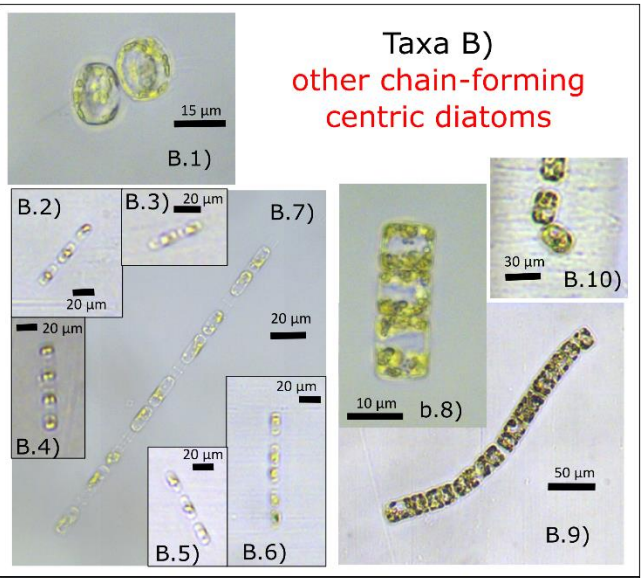
R^2 in %	C4	C3	C2
C1	45	11	1
C2	0	17	
C3	0		

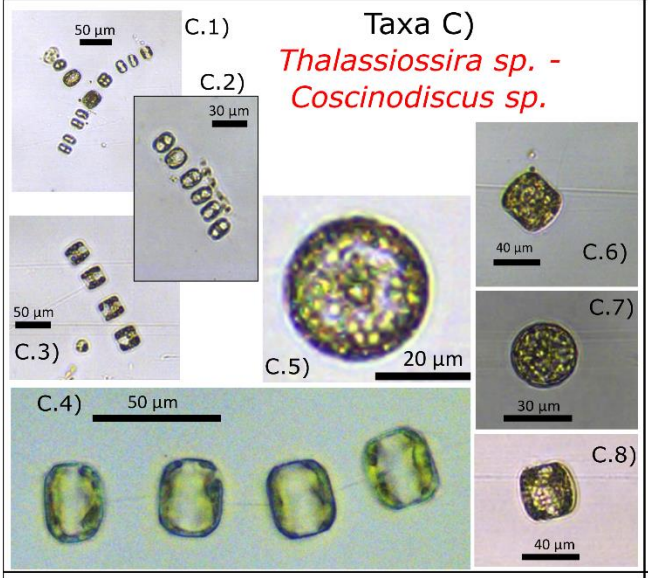
Appendix A: Properties and validation of a PARAFAC model with four components. Core consistency of 55% for 96.5% of variance explained. A) 3D representation of the associated data after variables' normalization (in a random sample taken for example). B) Scores and loadings plus leverage analysis. C) Split half validation figure (N=40 per group, best of 8 runs). D) Summary of components' correlation to each other, the value provided is an R^2 in percentage.

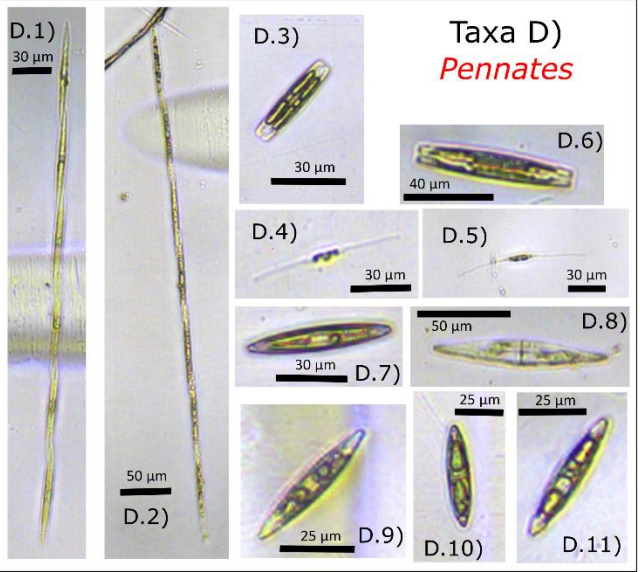


Appendix B: Fingerprints patterns. In a simplistic approach to periodicity, we hypothesize that typical minima and maxima observed from 200 to 450 Da could be explained using a 14 Da increment. To a certain extent, this was verified (pointed mass on the mass fingerprint) and different series starting at various points were identified for each component. Table B) summarizes these findings by identifying a series of maxima by red cells, and minima series by blue cells.









E)Population description

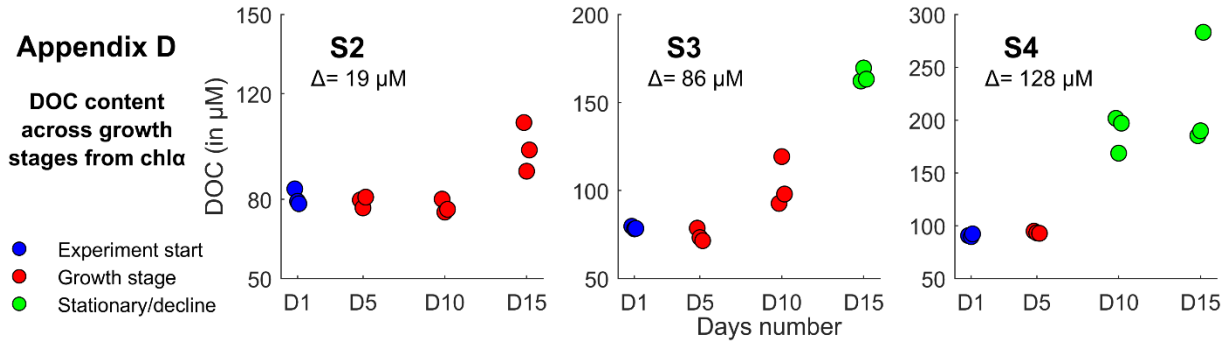
Category	*Range*	*Symbol*
Not observed	X	NO
rare	< 2%	- -
Few	2- 5%	-
Common	6- 10%	≈
Abundant	11- 25%	+
Dominant	> 25%	+ +

Qualitative population composition at D15

	Chaetoceros	Other chain-forming centric diatoms	Thal. - Cosc.	Pennates
S2-A	+ +	+ +	-	- -
S2-C	+	+ +	-	NO
S3-A	- -	+ +	≈	-
S3-B	- -	+ +	+	- -
S4-B	+	+ +	+	- -
S4-C	+ +	+ +	+ +	- -

Appendix C: microbial content analysis

An analysis of plankton composition was assessed over 2 of the three replicates of each "sizes-class" at each harvest. Diatoms were present in all, they were clustered in four taxonomic descriptors, which typical instances are here presented (A to D). A qualitative population description was also intended and the table E) present its results at Day 15.



***App. D*:** Change in DOC concentration across the experiment was monitored for each treatment. Dots' color refers to growth stages as determined earlier on using changes in chlorophyll α. Note the variable y-axis, such as the quantitative increase in µM from D1 to D15 was calculated (average of paired samples) and reported underneath each graph title.