

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

Chemical characterization of refractory dissolved organic matter

### Permalink

<https://escholarship.org/uc/item/2762h64d>

### Author

Arakawa, Neal Ken

### Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Chemical characterization of refractory dissolved organic matter

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy

in

Oceanography

by

Neal Ken Arakawa

Committee in charge:

Professor Lihini Aluwihare, Chair  
Professor William Gerwick  
Professor Chambers Hughes  
Professor Bradley Moore  
Professor Yitzhak Tor

2016

Copyright

Neal Ken Arakawa, 2016

All rights reserved

The Dissertation of Neal Ken Arakawa is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

---

---

---

---

---

Chair

University of California, San Diego

2016

## DEDICATION

This dissertation is dedicated to my family, friends, and previous academic mentors.

Thank you

## TABLE OF CONTENTS

Signature Page.....	iii
Dedication.....	iv
Table of Contents.....	v
List of Figures.....	vi
List of Tables.....	ix
Acknowledgments.....	x
Vita.....	xi
Abstract of the Dissertation.....	xii
Chapter I Introduction.....	1
Chapter II Direct identification of diverse alicyclic terpenoids in Suwannee River Fulvic Acid.....	11
Chapter III Carotenoids are the likely precursor of a major fraction of dissolved organic matter.....	32
Chapter IV Oxidation of $\beta$ -carotene produces compounds that resemble dissolved Organic matter.....	62
Chapter V Chemical characterization of solid-phase extracted dissolved organic matter across a salinity transect.....	87
Chapter VI Conclusions.....	128

## LIST OF FIGURES

Figure 2.1 A GCxGC selected-ion chromatogram (71 m/z) identifying the elution geography of n-alkane standards.....	15
Figure 2.2-2.5 GC FID chromatogram (2a) of reduced SRFA.....	16
Figure 2.6 95 m/z selected-ion chromatogram of reduced bald-cypress leaf litter.....	18
Figure 2.S1 Chemical reduction mechanism as published in Nimmagadda and McRae (I).....	21
Figure 2.S2 Overlaid GC-FID chromatograms of concurrent reductions with varying catalyst concentrations.....	26
Figure 2.S3 GC-MS spectra of peaks 1, 2, and 3 from Fig. S.2.....	27
Figure 2.S4 $^1\text{H}$ NMR of low catalyst reduction mixture (blue) and high catalyst reduction mixture (red).....	28
Figure 2.S5 Heteronuclear Single Quantum Coherence Spectroscopy (HSQC) spectrum of low level catalyst reduction mixture.....	29
Figure 2.S6 GC-MS spectrum of catalyst degradation peak.....	30
Figure 3.1 $^1\text{H}$ NMR spectra of 7 PPL-DOM, time series samples collected at the SIO Pier.....	49
Figure 3.2 Structures given for model compounds (Model A, Model B), and resulting reduced hydrocarbons (Comp A, Comp B).....	50
Figure 3.3 Structures for known carotenoid degradation products Loliolide and Absicic acid, and potential resulting reduced hydrocarbons.....	51
Figure 3.4 GCxGC selected-ion chromatogram (95 m/z) of chemically reduced, PPL-extracted marine DOM (PPL-DOM).....	52
Figure 3.5 Identification of intact carotenoids through $^1\text{H}$ - $^{13}\text{C}$ NMR correlations. One bond (A), and multiple bond (B) modelled $^1\text{H}$ - $^{13}\text{C}$ correlations for an intact carotene.....	53

Figure 3.6. (A). Long range $^1\text{H}$ - $^{13}\text{C}$ correlations for carotene. Region I shows terminal ring Alkane correlations, CDB and II show correlations to conjugated double bonds (orange in structure).....	54
Figure 3.7 Identification of oxidized carotenoids through $^1\text{H}$ - $^{13}\text{C}$ NMR correlations.....	55
Figure 3.8. $^1\text{H}$ NMR spectra of PPL-DOM (Red), Carotenoid-rich fraction (Green; retained on 2 <sup>nd</sup> PPL column following acid hydrolysis, see text), and sugar and amino acid-rich fraction (Blue; not retained by 2 <sup>nd</sup> PPL column following acid hydrolysis).....	56
Figure 3.9 Conceptual figure depicting how carotenoid degradation contributes soluble, structurally diverse molecules to the DOM reservoir. Photochemistry (hv) and other oxidative processes ( $\text{O}_2$ ) are depicted as leading mechanisms of post-biosynthesis modification.....	57
Figure 4.1 $^1\text{H}$ NMR spectra of $\beta$ -carotene (A), and fractions of oxidized $\beta$ -carotene products Residual starting material (B) Emulsion (C) and PPL Extractable (D)..	75
Figure 4.2. GC FID chromatograms of derivatized $\beta$ -carotene (A), and oxidized $\beta$ -carotene products.....	76
Figure 4.3 Observation of known $\beta$ -carotene degradation products by GC-MS.....	77
Figure 4.4 GCxGC-TOF-MS total ion chromatograms (TIC) of derivatized $\beta$ -carotene (A), and oxidized $\beta$ -carotene products Residual starting material (B) Emulsion (C) and PPL Extractable (D).....	78
Figure 4.5 Mass spectra from PPL Extractable and SIO Pier.....	79
Figure 4.6 $^1\text{H}$ NMR spectra of “Dark” (42d) control, purple, and light PPL extractable fractions: 7d, green; 24d, yellow; 42d, light blue, dark blue.....	80
Figure 4.7 $^1\text{H}$ NMR of filtered Seawater PPL Extractable (blue) vs MilliQ PPL Extractable (Red).....	81
Figure 4.8 95 m/z of reduction products of PPL extractable $\beta$ -carotene degradation products, and SIO Pier.....	82
Figure 4.9. Mass spectrum of compounds identified in reduced PPL Extractable products.....	83



Figure 4.10. Peak apex (spectrum at point in chromatogram, not background subtracted) of unresolvable regions in SIO Pier (1) and PPL Extractable (2) with mass spectrum which demonstrate strong similarities in reduction compounds.....	84
Figure 5.1. Sampling region near Skidaway Institute of Oceanography, GA.....	112
Figure 5.2. Elemental analyzer isotope ratio mass spectrometry (EA-IRMS) bulk data of sample set.....	113
Figure. 5.3 Representative GC-FID chromatograms of sample set. C11 (top) is a high salinity sample, and A4 (bottom) is a low salinity sample.....	114
Figure 5.4. FID integrations of Small molecule and UCM regions for bulk and hydrolyzed samples, (A, top), binned into salinity regions : “Low,” 0-10; “Mid,” 10-30; “High,” >30.....	115
Figure 5.5. GC-TOF-MS 73 m/z chromatograms for A6 bulk derivatized sample (top, A), and A6 hydrolyzed derivatized sample (bottom, B).....	116
Figure 5.6. Total area of all peaks integrated with 73 m/z ion.....	117
Figure 5.7. Total peak areas / mg sample as partitioned into classification regions.....	118
Figure 5.8. Total peak areas for samples as partitioned into classification regions.....	119
Figure 5.9. Evidence for Aminocaproic acid in samples.....	120
Figure 5.10. 204 m/z chromatogram of hydrolyzed C5 (High) sample , indicating sugar and sugar-like peaks in chromatogram.....	121
Figure 5.11. 95 m/z chromatograms of reduces SRFA (A), SIO-SPE (B), Low samples (C, E), High samples (D, F), and Mid sample (G).....	122
Figure 5.12 95 m/z 1D chromatogram of reduced samples shown in Fig. 5.11.....	123

## LIST OF TABLES

Table 2.1 Reduction of model compounds with n-Butylsilane/B(C <sub>6</sub> F <sub>5</sub> ) <sub>3</sub> .....	13
Table 3.1 Isotope and elemental data for PPL-DOM samples collected from SIO Pier (32.87°N, 117.26°W).....	58
Table 4.1. EA-IRMS data of β –carotene (A), oxidized β –carotene products.....	85
Table 5.1. List of study samples.....	124

## ACKNOWLEDGEMENTS

There are far too many people to thank for helping me throughout this dissertation. The most important are my advisor, Professor Lihini Aluiwhare, my friends, and my family.

I would also like to acknowledge the 2010 cohort and the Aluiwhare lab members for all of the good times over the years.

Chapter II, in full, is a reprint of the material as it appears in Direct identification of diverse alicyclic terpenoids in Suwannee River Fulvic Acid in *Environmental Science and Technology*, 2015, Arakawa, Neal; Aluiwhare, Lihini, 49: 4097-4105. The dissertation author was the primary investigator and author of this paper.

Chapter III, in part, has been submitted for publication of the material as it may appear in *Science Advances*, 2016, Arakawa, Neal; Lane-Coplen, Daniel; Soong, Ronald; Stephens, Brandon; Simpson, Andre J.; Aluiwhare, Lihini. The dissertation author was the primary investigator and author of this paper.

Chapter IV, in part is currently being prepared for submission for publication of the material. Arakawa, Neal; Aluiwhare, Lihini. The dissertation author was the primary investigator and author of this material.

## VITA

- 2016            Doctor of Philosophy, Scripps Institution of Oceanography
- 2010-2016     Research Assistant, Scripps Institution of Oceanography
- 2008            Bachelor of Science, University of California, San Diego

## PUBLICATIONS

Direct identification of diverse alicyclic terpenoids in Suwannee River Fulvic Acid.  
Environmental Science & Technology 49: 4097-4105.

## ABSTRACT OF THE DISSERTATION

Chemical characterization of refractory dissolved organic matter

by

Neal Ken Arakawa

Doctor of Philosophy in Oceanography

University of California, San Diego, 2016

Professor Lihini I. Aluwihare, Chair

The primary objective of this thesis was to combine a chemical degradation technique together with an analytical framework centered primarily around gas chromatography (GC) to more fully interrogate the composition of aquatic dissolved organic matter (DOM). Previous studies had suggested that aliphatic compounds could

represent a significant fraction of refractory organic matter isolated by solid phase extraction (SPE). These studies had also uncovered the vast complexity of DOM. Gas chromatography coupled to mass spectrometry provides superior separation capability and is ideal for examining complex mixtures of lipid-derived molecules. As such I sought to develop a comprehensive GC analysis methods to provide molecular level information for DOM isolated by solid phase extraction (SPE) onto a hydrophobic resin- PPL (Agilent Bond Elut). In Chapter II, a comprehensive chemical reduction procedure was developed and first applied to the environmental DOM standard Suwannee River Fulvic Acid (SRFA) as a proxy for marine DOM. The resulting hydrocarbons were amenable to comprehensive gas chromatography time-of-flight mass spectrometry (GCxGC-TOF-MS), and effectively resolved into multiple series of alicyclic, unsaturated compounds. This was the first direct demonstration of the isomeric complexity of aquatic DOM. Similar alicyclic compounds were recovered from the reduction of terrestrial source material, implicating resin acids and sterols as potential precursors of SRFA. In Chapter III the reduction process was applied to marine surface DOM from the Scripps Institution of Oceanography Pier, and similar alicyclic compounds were found. The GCxGC-TOF-MS identified carbon backbones closely resembling carotenoids, implicating these ubiquitous and highly reactive biomolecules as the source of a significant fraction of DOM accumulating in the marine water column. The structural assignment was supported by the identification of carotenoid derived resonances in two dimensional nuclear magnetic resonance (NMR) spectra, which indicated that these molecules were highly oxidized compared to the parent molecules consistent with their present in DOM. Following up on this work in Chapter IV the carotenoid  $\beta$ -carotene was irradiated with

natural sunlight to test the hypothesis that photodegradation was one pathway that converted carotenoids into water-soluble degradation products. The first finding was that the reaction produced a series of compounds identical to compounds isolated from marine DOM. The second important result was that the reaction produced a complex mixture of isomers from a single compound that helps to at least partly explain the compositional diversity in marine DOM. Together, the data in Chapters III and IV allowed us to link a large fraction of DOM to a ubiquitous biomolecule that can now serve as a model for studies examining the formation and fate of DOM that accumulates in the ocean on long timescales. Finally, in Chapter V we sought to examine how the composition of DOM – both the complex alicyclic fraction and small, polar biomolecules, which are considered a “fresher” signal of biological input – evolved across a salinity gradient. Although core biochemical classes were present in all regions the data supported *in situ* production of compositionally similar material rather than mixing across the gradients as proposed in some studies. Together, the chapters in my thesis provide new insight in the composition of dissolved organic matter in marine and terrestrial environments. The thesis also represents the most comprehensive molecular level characterization of DOM isolated by this solid phase extraction method, which is the most commonly used isolation method in the field. My findings also provide an important foundation for future lab-based mechanistic studies of DOM cycling in the marine environment.

## **Chapter I**

### Introduction



### *Dissolved organic matter (DOM) and Dissolved organic carbon (DOC)*

Dissolved organic matter (DOM) and dissolved organic carbon (DOC) are two terms that are used, often interchangeably, to describe compounds that are dissolved in natural waters (e.g. rivers, lakes, oceans). As consistently as possible, I will use DOM when discussing qualities of entire compounds (such as chemical structure) and DOC when discussing qualities of the carbon contained within those compounds (such as concentration or radiocarbon age). DOM is an operational definition; compounds are referred to as DOM (and DOC) *if* they pass through a filter and are *not* inorganic (e.g.  $\text{CO}_2$ ,  $\text{HCO}_3^-$ ,  $\text{CO}_3^{2-}$ ). The pore size of this filter is not standardized, and ranges from 0.1-1 $\mu\text{m}$ . In Chapters 3 and 4 I use a 0.2 $\mu\text{m}$  filter, which largely excludes bacteria.

### *Global significance of DOC*

As a reservoir, global oceanic DOC is estimated at 662 Gt (1Gt =  $10^{15}$ g) of carbon (C) (Hansell et. al. 2009). It is similar in scale to the amount of carbon contained in the atmosphere as  $\text{CO}_2$ , ~760 Gt C, or the amount of carbon contained within the terrestrial biosphere, 800 Gt C (Falkowski et al. 2000). While terrestrial carbon is largely contained within woody tissue (and thus temporarily protected from oxidation), aquatic DOC is easily accessible to heterotrophic bacteria and photooxidation. Given that the ocean surface is saturated in oxygen and light can penetrate even 100s of meters in most open ocean environments, I would expect organic compounds to be oxidized completely to  $\text{CO}_2$ . However, early radiocarbon measurements ( $^{14}\text{C}$ ) (Druffel, 1987, Druffel et al 1992) indicated that deep ocean DOC has an average radiocarbon age of 4000-6000 years, with some compounds perhaps being significantly older; recent studies classifying DOM into

different reactivity states estimate that 95% of oceanic DOC is not easily accessible to bacterial degradation or photooxidation (Hansell 2013). Since much of the deep ocean reservoir of DOC has a radiocarbon age that exceeds the mixing time of the ocean, old, degradation resistant organic compounds are mixed throughout the entire oceanic water column. Given the capacity of DOC to effectively double atmospheric CO<sub>2</sub>, the cause of this recalcitrance is the key to understanding the full role of DOC in impacting past (Sexton, 2011), as well as future, climates.

#### *Chemical characterization of DOM*

Photosynthetically produced biochemicals (e.g. sugars, amino acids, lipids) are the source of most of the carbon in DOM, with small contributions from chemoautotrophy and hydrothermal systems. Biologically mediated and abiotic processes transform these compounds into DOM, such that known biochemicals represent less than 10% at the surface and less than 2% in deep water (Benner, 2002). The majority of DOM therefore exists as a chemically complex mixture of innumerable compounds. Most chemical studies of DOM have utilized isolation methods that fractionate DOM by size or chemistry, and none of these methods can isolate >60% of DOC on their own. Currently under development is a method that couples reverse osmosis and electro dialysis, which is capable of isolating nearly the entire reservoir. However, it is currently difficult to implement and restricted to only a few laboratories (Koprivnjak et al. 2009). Largely, studies characterizing DOM have used one of two general techniques to isolate DOM: ultrafiltration (UF) or solid-phase extraction (SPE). As each targets a different attribute of DOM (UF: size, SPE: hydrophobicity) with

different efficiencies, these methods introduce a sampling bias that must be acknowledged. In these studies (Chapters 3-5) analyze DOM isolated by SPE using a proprietary resin identified as PPL (Agilent Bond-Elut). This particular SPE method consistently isolates a larger (%) fraction – 40-60% - of available DOC. Its radiocarbon signature is as depleted as bulk DOC (-180-200‰) at least in surface waters (Lechtenfeld, 2014) whereas ultrafiltration consistently isolates a fraction that is younger than the bulk (Aluwihare et al. 2002) and other SPE resins such as XAD, isolates a more depleted fraction (e.g., Druffel et al., 1992).

Chemical analyses of DOM range from bulk to highly targeted, with each offering distinct advantages and limitations. One such bulk measurement is the isotopic ratio of naturally occurring  $^{13}\text{C}/^{12}\text{C}$  within DOM, which has long been used to distinguish marine from terrestrial DOM. Bulk measurements are attractive because they integrate the entire sample, and does not introduce an additional bias. However, chemical characterization *is* limited by analytical bias. Unfortunately, even high-powered bulk techniques such as two-dimensional nuclear magnetic resonance spectroscopy (e.g., HSQC, HMBC) ultimately reach analytical limitations due to the overwhelming complexity of DOM.

Recently, ultrahigh resolution mass spectrometry (FTICR-MS) has emerged as perhaps the most popular analytical tool for DOM analysis. FTICR-MS is a prime example of the power and limitations of targeted analytical techniques. Functionally, FTICR-MS resolves complexity by separating DOM ions in a strong magnetic field. The strength of this separation (or resolving power), and accuracy of the instrument are sufficiently high enough to determine exact chemical formulae for many of the ions, an

unprecedented feat in DOM analysis. While inferences about chemical composition can be drawn based on the elemental ratios of constructed formulas, FTICRMS does not provide structural information. Additionally, all mass spectrometric techniques require ionization, which also introduces bias to the analysis.

This discussion underscores the importance of combining and integrating different analytical approaches when analyzing DOM structure. Arguably, the most important “breakthroughs” in DOM structure elucidation have occurred by combining analytical techniques. Recently, FTICR-MS and multiple NMR techniques (i.e. HSQC, DEPT  $^{13}\text{C}$ ) were combined to model a component of DOM present in deep (old) DOM (Hertkorn et al., 2006). Structural features present in these NMR spectra gave birth to the term carboxy-rich alicyclic molecules (CRAM) to describe this fraction of DOM. These features are consistent with carboxylated sterol structures, however, CRAM cannot be definitively tied to any particular class of compounds, and uncovering the connectivity between the prominent functional groups to provide plausible chemical structures remains a challenge. More recently, a distinct class of oxygen-rich, aliphatic compounds was identified in terrestrial DOM, through the combination of HSQC and HMBC NMR analysis (Lam et al., 2007). This class was labeled molecules resembling degraded, linear, terpenoids (MDLT), and is distinct from CRAM.

A central theme in this dissertation is the use of chemical alteration of DOM in combination with instrumental analysis. Chemical degradation/alteration of DOM is not novel and has historically been applied to combat the complexity of the DOM. For example, cupric oxide ( $\text{CuO}$ ) based oxidation of organic matter to analyze lignin

oxidation products was first applied to organic matter samples 40 years ago (Hedges and Parker 1976), and is still widely used today. Other popular techniques include TMAH thermochemolysis, acid hydrolysis (Panagiotopoulos et. al. 2007), and recently benzene polycarboxylic acid (BPCA) analysis (Glaser et. al. 1998). In fact, exhaustive chemical reduction of marine dissolved organic matter was also accomplished ~40 years ago (Gagosian and Sturmer, 1977). Had currently available high-resolution analytical methods, and specifically, superior chromatographic separation, been accessible at the time, these authors would have likely reached the same conclusions as I am present in my thesis. The structural characterization that I performed utilized comprehensive gas chromatography time of flight mass spectrometry (GCxGC-TOF-MS). This instrument greatly increased our separation power, and as a result, our ability to provide definitive chemical structures for some fraction of PPL extracted DOM.

GCxGC-TOF-MS is a powerful tool which increases gas chromatographic resolution through the inclusion of a second, in-line, column. While its most direct application is the separation of oil samples (as no derivatization is necessary), it can be applied to any extremely complex mixture of compounds amenable to gas chromatography, and has already shown promise in analyzing environmental samples (Ball & Aluwihare, 2014).

### *Organization of the dissertation*

The central goal of this dissertation is to better understand the chemical composition of DOM, with a particular focus on the aliphatic compounds that accumulate in this reservoir. All of the chapters include the use of an exhaustive chemical reduction

method to make DOM amenable to comprehensive GC (GCxGC) TOF-MS.

Comprehensive GC affords remarkable chromatographic resolution, making it perhaps the most powerful tool currently available for structural elucidation of complex mixtures. When possible, we have complimented chemical reduction data with other analytical datasets to lend further support to our conclusions.

Chapter II of this dissertation applied a previously published, chemically exhaustive reduction method to an environmental DOM standard, Suwannee River Fulvic Acid (SRFA). This chapter was published in *Environmental Science and Technology* and introduced a novel analytical framework – coupling DOM reduction to comprehensive GC TOF-MS – to provide fundamentally new data on the chemical backbones present in DOM. Ultimately, the study identified alicyclic hydrocarbons as prominent reduction products of terrestrial DOM (Arakawa and Aluwihare, 2015). The study also applied the reduction to a broad suite of standard molecules, which provided a template for proposing parent structures for reduction products from complex environmental samples.

Chapter III applies the reduction method developed in Chapter II to a marine DOM sample isolated from the Scripps Institution of Oceanography (SIO) Pier by SPE. This chapter is currently in review at *Science Advances*. In this chapter we determine that alicyclic reduction products are present in  $^{14}\text{C}$  aged SIO-SPE DOM, and that these hydrocarbons originate from degraded carotenoids. This work is supported by advanced NMR experiments conducted in collaboration with colleagues at the University of Tortona, Canada. The reduction coupled to comprehensive GC TOF-MS uncovered

molecules with carotenoid-like carbon backbones and the NMR experiments successfully tied these compounds to oxidized carotenoid-structures accumulating in DOM.

Chapter IV tests the hypothesis that carotenoids can be degraded to form compounds which resemble DOM. In this chapter a single carotenoid,  $\beta$ -carotene, was oxidized over relatively short (~1 month) time periods. Using standard compounds, comprehensive GC, and proton NMR experiments I was able to conclusively demonstrate the structural similarities between the diverse set of carotenoid photo-oxidation products and PPL-DOM from the SIO Pier.

Chapter V examines a suite of SPE-DOM samples that span a salinity transect from river to ocean. This chapter uses a variety of bulk and targeted techniques to differentiate SPE-DOM across the gradient, and connect individual molecular features, if possible, to DOM produced in different salinity regimes. The data suggests that while certain bulk measurements support mixing between high and low salinity endmembers as the major driver of compositional differences, the molecule techniques (including chemical reduction) point to three distinct compositional pools of SPE-DOM.

## References

- Aluwihare, L.I., Repeta, D.J., Chen, R.F. 2002. Chemical composition and cycling of dissolved organic matter in the Mid-Atlantic Bight. *Deep Sea Res. II* 49: 4421-4437
- Arakawa, N., Aluwihare. 2015. Direct Identification of diverse alicyclic terpenoids in Suwannee River Fulvic Acid. *Environmental Science and Technology* 49: 4097-4105.
- Ball, G.I., Aluwihare, L.I. 2014. CuO-oxidized dissolved organic matter (DOM) elucidated by comprehensive two dimensional gas chromatography-time of flight-mass spectrometry (GCxGC-TOF-MS). *Organic Geochemistry* 75:87-98
- Benner, R. 2002. Chemical composition and reactivity in Biogeochemistry of Marine Dissolved Organic Matter. Hansell, D.A., Carlson, C.A., eds. Academic Press: 59-90
- Druffel, E.R.M., Williams, P.M. 1987. Radiocarbon in dissolved organic matter in the central North Pacific Ocean. *Nature* 330: 246-248.
- Druffel, E.R.M., Williams, P.M., Bauer, J.E., Ertel, J.R., 1992. Cycling of dissolved and particulate organic matter in the open ocean. *Journal of Geophysical Research* 97: 15,639-15,659
- Falkowski, P., Scholes, R., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Hogberg, P., Linder, S., Mackenzie, F.T., Moore, B., Pederson, T., Rosenthal, Y., Seitzinger, S., Smatacek, V., Steffen, W. 2000. The global carbon cycle: a test of our knowledge of earth as a system. *Science* 290:291-296.
- Gagosian, R.B., Stuermer, D.H. 1977. The cycling of biogenic compounds and their diagenetically transformed products in seawater. *Marine Chemistry* 5: 605-632
- Glaser, B. Haumaier, L., Guggenberger, G. Zech, W. 1998. Black carbon in soils: the use of benzenecarboxylic acids as specific markers. *Organic Geochemistry* 29: 811-819.
- Hansell, D.A., Carlson, C.A., Repeta, D.J., Schlitzer, R. 2009. Dissolved organic matter in the ocean: a controversy stimulates new insights. *Oceanography* 22: 202-211.



- Hansell, D.A. 2013. Recalcitrant dissolved organic carbon fractions. *Annual review of marine science* 5: 421-445.
- Hedges, J., Parker, P. 1976. Land-derived organic matter in surface sediments from the Gulf of Mexico. *Geochim Cosmochim Acta* 40:1019-1029
- Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., Kettrup, A., Hedges, J.I. 2006. Characterization of a major refractory component of marine dissolved organic matter. *Geochim Cosmochim Acta* 70: 2990-3010
- Koprivnjak, J., Pfromm, P.H., Ingall, E., Vetter, T.A., Schmitt-Kopplin, P., Hertkorn, N., Frommberger, M., Knicker, H., Perdue, E.M. 2009. Chemical and spectroscopic characterization of marine dissolved organic matter isolated using coupled reverse osmosis-electrodialysis. *Geochim Cosmochim Acta* 73: 4215-4231
- Lam, M., Baer, A., Alae, M., Lefebvre, B., Moser, A., Williams, A., Simpson, A.J. 2007. Major structural components in freshwater dissolved organic matter. *Environmental Science and Technology* 41: 9240-9247
- Lechtenfeld, O. J., Kattner, G., Flerus, R., McCallister, S. L., Schmitt-Kopplin, P., Koch, B. P. 2014. Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean. *Geochim. Cosmochim. Acta* 126: 321-327
- Panagiotopoulos, C., Repeta, D. J., Johnson, C.J. 2007. Characterization of methyl sugars, 3-deoxysugars and methyl deoxysugars in marine high molecular weight dissolved organic matter. *Organic Geochemistry* 38: 884-896
- Sexton, P.F., Norris, R.D., Wilson, P.A., Palike, H., Westerhold, T., Rohl, U., Bolton, C.T., Gibbs, S. 2011. Eocene global warming events driven by ventilation of oceanic dissolved organic carbon. *Nature* 471: 349-352.

## **Chapter II**

Direct Identification of Diverse Alicyclic Terpenoids in Suwannee River Fulvic Acid

## Direct Identification of Diverse Alicyclic Terpenoids in Suwannee River Fulvic Acid

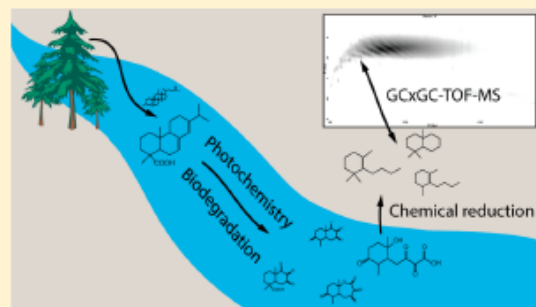
Neal Arakawa\* and Lihini Aluwihare

Scripps Institution of Oceanography, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093-0244, United States

### Supporting Information

**ABSTRACT:** The chemical complexity of dissolved organic matter (DOM) obstructs our ability to definitively recover source compounds from within DOM, an objective which has the capacity to alter our understanding of carbon sequestration on a global scale. To advance compositional studies of DOM we have applied a previously published reduction method to an environmental standard, Suwannee River Fulvic Acid (SRFA). The reduction products, comprising 12% of the pre-reduced carbon, were then separated by comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC×GC-TOF-MS). Results indicate that the majority of observed reduced compounds corresponded to alicyclic hydrocarbons in the size range C<sub>10</sub> to C<sub>17</sub>. Cyclic terpenoids are the only biomolecule class with contiguous, alicyclic carbon

backbones of this size. These terpenoid reduction products contain series offset by CH<sub>2</sub> and exhibit great isomeric diversity, features previously inferred from ultrahigh resolution mass spectrometry and NMR studies of unreduced SRFA. Reduction of *Taxodium* leaf litter as a source material to SRFA confirmed the prevalence of terpenoids in SRFA and provided insight into the parent compounds that must be diagenetically modified on relatively short time scales. These data corroborate several recent studies that suggest alicyclic hydrocarbons to be important components of longer-lived DOM.



### INTRODUCTION

Dissolved organic matter (DOM) represents a quantitatively significant reservoir of the global carbon cycle.<sup>1,2</sup> Yet, its provenance remains elusive, as much of it has been immune to degradative compositional interrogation. Recent studies employing nuclear magnetic resonance spectroscopy<sup>3–5</sup> (NMR) and Fourier transform ion cyclotron resonance spectrometry<sup>6,7</sup> (FT-ICR-MS) have implicated linear and alicyclic terpenoids as major precursors of terrestrial and marine DOM. Sources of terpenoids and their biosynthetic pathways are well-known, and so, direct identification of the relevant structures and their diversity could hold significant promise for linking DOM composition to source.<sup>8</sup>

Some of the earliest reports of terpenoid precursors in DOM are from studies of Suwannee River Fulvic Acid (SRFA), a widely used environmental DOM standard. The sum of the body of current work indicates that SRFA is derived from a mixture of lignin, tannin, and terpenoid (including sterol) precursors.<sup>3–5,9</sup> Yet, quantitatively little lignin, and no definitive tannin or terpenoid structures, have been isolated from humic substances.<sup>10</sup> The recovery of specific structures from SRFA and other natural organic matter (NOM) would not only confirm source material identity, but also has the potential to inform our understanding of the processes that alter DOM on the molecular level. As with previous studies we use SRFA as a

model for DOM and test a method that is specifically designed to access hydrocarbon backbone structures.

Although fractionation of SRFA via liquid chromatography is improving, isolation of individual compounds is currently unattainable.<sup>11</sup> Gas chromatography (GC) offers better separation, but much of SRFA is likely too large and polar to be analyzed natively.<sup>6,7</sup> However, chemical degradation, such as chemical reduction, is a viable method for making SRFA more GC-amenable. Previous studies employing an *n*-butylsilane reduction method reported the separation of various products from SRFA, but much of the mixture was left unresolved.<sup>12,13</sup> In this study, the reduction method has been adapted and combined with comprehensive gas chromatography (GC×GC) coupled to time-of-flight mass spectrometry (TOF-MS), to provide additional molecular-level data on the composition of SRFA. This GC×GC technique uses an additional column to provide unsurpassed chromatographic resolution, especially when applied to hydrocarbons.<sup>14</sup> Data presented here confirms the presence of cyclic terpenoid derived carbon backbones in SRFA. The unique capability of this approach complements the

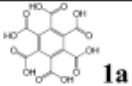
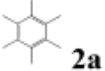
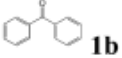
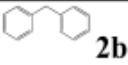
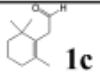

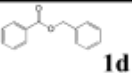

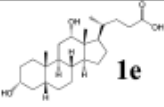
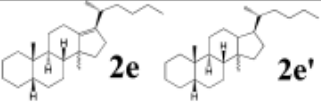

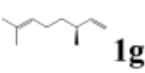
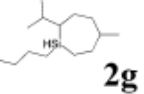

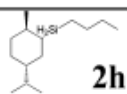
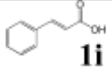

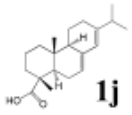
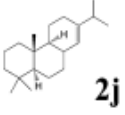
Received: November 11, 2014

Revised: March 10, 2015

Accepted: March 13, 2015

Published: March 13, 2015

Chart 1. Reduction of Model Compounds with *n*-Butylsilane/ $B(C_6F_5)_3$ <sup>a</sup>

Entry	Substrate (1)	Product(s) (2)	GC Yield
1	 <b>1a</b>	 <b>2a</b>	58%
2	 <b>1b</b>	 <b>2b</b>	78%
3	 <b>1c</b>	 <b>2c</b>	65%
4	 <b>1d</b>	 <b>2d</b>	65%
5	 <b>1e</b>	 <b>2e</b> <b>2e'</b>	72%
6	 <b>1f</b>	No observed product	N.A.
7	 <b>1g</b>	 <b>2g</b>	*25%
8	 <b>1h</b>	 <b>2h</b>	*25%
9	 <b>1i</b>	 <b>2i</b>	51%
10	 <b>1j</b>	 <b>2j</b>	34%

<sup>a</sup>Products 2g and 2h coelute. GC yield calculated as average.

existing suite of degradation techniques and can provide a more comprehensive molecular level understanding of DOM source and fate.

## MATERIALS AND METHODS

**Materials.** All solvents (ACS grade or better), model compounds and hydrocarbon standards were purchased commercially, and used as supplied. Suwannee River Fulvic Acid (2S101F) was purchased from the International Humic Substances Society. *n*-Butylsilane (*n*-BS) and Tris-

(pentafluorophenyl)borane ( $B(C_6F_5)_3$ ) were purchased from Sigma-Aldrich, and stored in an argon filled desiccator when not in use. Dichloromethane ( $CH_2Cl_2$ ) was stored over activated molecular sieves (3 Å, 10% wv). Freshly fallen leaf litter (based on intactness of leaves) was collected from Tifton, GA, and verified to be of the genus *Taxodium*. Leaves were freeze-dried for 1 week, and ground using a mortar and pestle before reduction.

**Reduction Procedure.** The reduction procedure was modeled after previous studies.<sup>12</sup> For model compound

reductions, compounds (~1 mg of each) were weighed (balance accurate to 0.1 mg) into a single, flame-dried 2 mL vial equipped with a stir bar. Under argon atmosphere, 10 mg of  $B(C_6F_5)_3$  was added (100  $\mu$ L of 100 mg  $B(C_6F_5)_3$ /mL  $CH_2Cl_2$  solution). Immediately after addition, 100  $\mu$ L of n-BS was added, also under argon. The mixture was then allowed to stir overnight. Following reduction, the samples were treated by the following procedure to remove excess catalyst and siloxane byproducts. Subsequently 50  $\mu$ L of the pot was syringed into 250  $\mu$ L of conc. potassium hydroxide in methanol (KOH/MeOH). After 1 min, the mixture was extracted with pentane ( $3 \times 100 \mu$ L). The pentane fractions were collected, washed with  $H_2O$ , and dried over sodium sulfate. The organic fraction was then dried under  $N_2$  to 100  $\mu$ L for GC-FID/MS and GC $\times$ GC-TOF-MS analysis. The same procedure was used for the reduction of SRFA and leaf litter, with 10 mg of starting material for each. All reductions were carried out in duplicate. While SRFA was reduced to dissolution, particulates were still observed in the leaf litter reduction.

Reaction conditions were selected based on results from model compound reductions. The amount of  $B(C_6F_5)_3$  used was found to be the most critical parameter affecting the reduction yield of model compounds. Chart 1 lists model compounds and their observed reduction products. In previous work with **1a** (10 mg), which was the most chemically resistant of the model compounds tested, it was observed that increasing the amount of  $B(C_6F_5)_3$  to 10 mg also increased reduction yields, whereas increasing n-BS did not. The ratio of 10 mg  $B(C_6F_5)_3$  to 10 mg starting material was therefore selected to ensure complete reduction of oxygen containing groups in model compounds as well as in environmental samples. This quantity of catalyst gave the most predictable and consistent results across all model compounds and SRFA.

**Gas Chromatography-Flame Ionization Detection/Mass Spectrometry (GC-MS/FID).** An Agilent 7890A gas chromatograph coupled to an Agilent 5975C quadrupole mass spectrometer was used for chromatographic separation and compounds were detected simultaneously with Electron Ionization MS and FID, enabled by a splitter. Splitless injection with 1  $\mu$ L of analyte was performed onto a 5% phenyl poly(dimethylsiloxane) column (J&W DB-ht, 30 m, 320  $\mu$ m i.d., 0.1  $\mu$ m film) and separated using a temperature program from 50  $^\circ$ C (hold time 0.5 min) to 140  $^\circ$ C (at 10  $^\circ$ C/min, hold time 0 min) to 320  $^\circ$ C (at 15  $^\circ$ C/min, hold time 10 min). Helium was the carrier gas with a constant flow of 1.8 mL/min, the FID was operated at 310  $^\circ$ C and the mass spectrometer conditions included 70 eV ionization with scanning from 50 to 750  $m/z$ . The transfer line was set at 280  $^\circ$ C. Quantification of reduced model compounds and reduced SRFA was accomplished using a hydrocarbon standard (decahydronaphthalene) and assuming a constant FID response factor across hydrocarbons. When available, reduced model compound structures were identified by comparison to the NIST Mass Spectral Standard Reference Database (2005). For SRFA, manual integration was used to remove peaks associated with catalyst degradation and residual siloxanes, and sample MS are located in the Supporting Information.

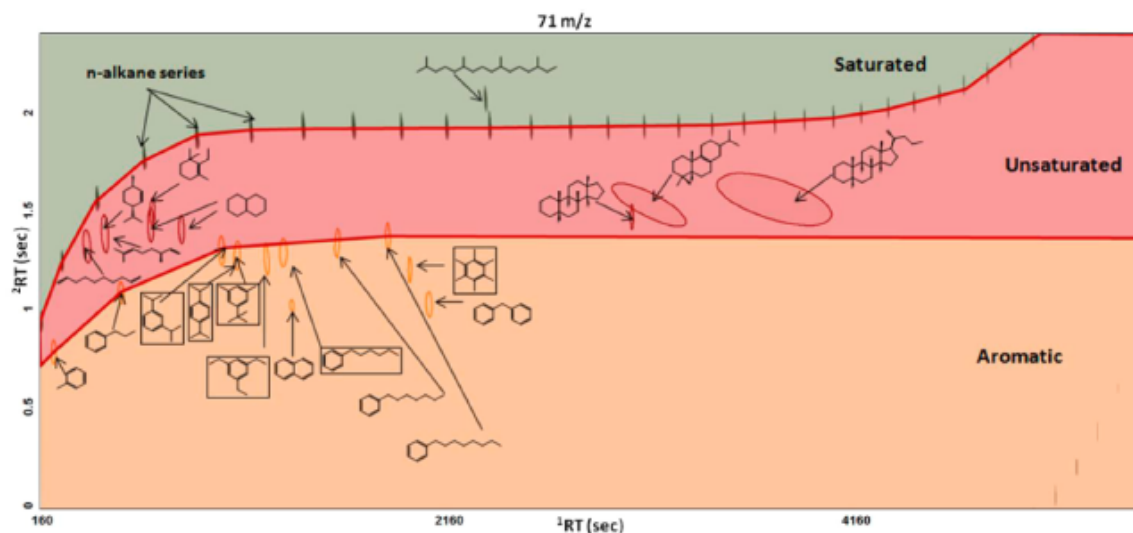
**Comprehensive Two-Dimensional Gas Chromatography (GC $\times$ GC-TOF-MS).** A LECO Pegasus 4D GC $\times$ GC-time-of-flight mass spectrometer (GC $\times$ GC-TOF-MS) was used for comprehensive GC analysis.<sup>15</sup> The term GC $\times$ GC refers to the use of two orthogonal columns which were selected to increase the separation of aliphatic from aromatic hydrocarbons. The

analysis is comprehensive because all of the effluent from the first column is cryofocused and transferred onto the second column. Effluent from the second column is then analyzed by the TOF-MS, which benefits from high spectral acquisition speed (50 Hz–500 Hz). Finally, the data is compiled into a two-dimensional chromatogram that is visualized and processed by ChromaTOF software. Both columns are housed within an Agilent 7890A gas chromatograph. The splitless inlet temperature was set at 300  $^\circ$ C. The first dimension column was a semipolar Crossbond diphenyl dimethyl polysiloxane column (Restek Rxi-17Si, 30 m length, 0.255 mm i.d., 0.25  $\mu$ m film thickness). The column was programmed to remain isothermal at 40  $^\circ$ C for 1 min, and ramped to 315  $^\circ$ C at 3  $^\circ$ C per minute. The modulator temperature was offset by +30  $^\circ$ C to the primary oven. The secondary oven (within the GC) housed the second dimension non polar Crossbond dimethyl polysiloxane column (Restek Rxi-1, 1.58m length, 0.250 mm i.d., 0.25  $\mu$ m film thickness). The secondary oven temperature was offset by +25  $^\circ$ C to the primary oven. The modulation period was 2.5 s, with a hot pulse time of 1.05 s and a cool time of 0.2 s. The carrier gas was Helium at a constant flow of 1.5 mL/min. The acquisition delay on the TOF-MS was set to 160 s. The acquisition rate was set to 50 Hz, with a range of 5–1000  $m/z$ . The transfer line was set at 280  $^\circ$ C. Electron Impact was run at 70 eV. Again, reduced model compound structures were identified by comparison to the NIST Mass Spectral Standard Reference Database (2011).

## RESULTS AND DISCUSSION

**Model Compound Analysis.** The reduction process was applied to a collection of chemically distinct model compounds with functional groups that are expected to also be present in SRFA. This was done to verify that the products were both consistent with previously published literature and consistently reproducible. We observed that the catalyst ( $B(C_6F_5)_3$ ) and byproduct siloxane polymers often coelute in high concentrations with analytes of interest, impacting accurate GC-FID quantification and complicating GC $\times$ GC analysis. We used concentrated KOH/MeOH to remove these byproducts, which resulted in cleaner chromatograms. However, this method also removes reduction products which are polysilylated, such as phenols (the reduction products of lignins and tannins) and some hydrosilylated (e.g., **2g**, **2h**) compounds. Recovery of phenols can be achieved by acid hydrolysis of silyl ethers, but are not the topic of this study. Below, we systematically describe the effect of chemical reduction on different functional groups, as supported by selected substrates and their reduced products.

The reduction process was adapted from previous work,<sup>12,13</sup> and the reaction mechanism is reproduced in the Supporting Information. All of the carboxylic acids in model compounds (**1a**, **1e**, **1i**, and **1j**) reduced completely to methyl groups, as predicted. The structure of **1a** is likely to rival the most sterically hindered arrangement of oxygen containing functional groups in SRFA, and is potentially a good chemical proxy. Ketones such as in **1b** are labile, reducing completely to methylene groups. Aldehydes such as in **1c** are also labile, reducing completely to methyl groups. Ester linkages, such as in **1d**, cleaved and reduced completely. This product was not expected based on previous studies<sup>12</sup> in which reduction of the carbonyl but preservation of the newly created ether linkage was observed. Aliphatic alcohols are reduced, but as is shown in



**Figure 1.** A GC×GC selected-ion chromatogram (71  $m/z$ ) identifying the elution geography of  $n$ -alkane standards. Circled regions overlay areas where depicted reduced model compounds and hydrocarbons elute. Boxed compounds are isomers of aromatic  $C_{12}H_{18}$  compounds. Note that the reduction products of **1e** and **1j** elute as a series of peaks over a large region.

**1e** they can also be eliminated in particular cases, resulting in a double bond.

In contrast to previous studies, we have found that with a few exceptions double bonds are labile, and this can be attributed to the higher catalyst concentrations employed here. In the reduction of linear diene **1f** no products are observed. This is believed to be the result of a hydrosilylation reaction at the double bond(s). The resulting organosilane product further reacted with reducing agent byproducts, forming compounds which are lost during the postreduction workup. This behavior is observed in branched diene **1g** and cyclic alkene **1h**. The hydrosilylation products **2g** and **2h** are in fact isolated but in low yields, which indicates that material is lost to further reactions. Hydrosilylation of double bonds has been observed in previous literature under similar conditions.<sup>16</sup> The incorporation of  $n$ -BS into products is easily identified by distinctive fragmentation patterns in mass spectra.

In addition to hydrosilylation, double bonds can also be reduced under reaction conditions. In the reduction of **1i**, the isolated product **2i** is fully reduced. However, the reduction product **2j** is only partially reduced. This is because double bonds present in a ring system are more resistant to reduction, as is observed for products **2c** and **2e**. That double bonds are most stable between tertiary carbons is clearly seen in the difference between the reduction products **2c** and **2h**, where the double bond is preserved in **2c** but hydrosilylated in **2h**. The mechanism by which double bonds are reduced vs hydrosilylated is not yet fully understood. There does appear to be a correlation of reduction with conjugated compounds (**1i**, **1j**), however in a later section we observe that nonconjugated terpenes are also reduced.

The reduction of **1e** and **1j** produces multiple isomers of products with identical molecular ions and similar fragmentation patterns. Previous studies have shown that Lewis acid catalysts enable backbone rearrangement (double bond migration and methyl shifts) in unsaturated sterol hydrocarbons.<sup>17</sup> In those studies and our current work, rearrange-

ments give rise to multiple isomers from the reduction of a single compound, and the relative amounts of these products can be altered by the amount of catalyst used. A combination of GC-MS,  $^{13}C$  NMR, and  $^1H$  NMR of the reduction products was used to determine the location of the double bond in **2e**. A detailed description of this derivation is provided in the Supporting Information. Thus, far, isomer production has only been observed in fused cyclic structures, which contained either double bonds or hydroxyl groups directly bonded to the ring system prior to reduction. It is further proposed, though not tested here, that other functional groups incorporated into the ring backbone, such as ketones, would also behave in a similar manner to the hydroxyl reduction.

Finally, we have observed that the reduction of **1e** produces both fully (**2e'**) and partially reduced products (**2e**), with a molecular ion offset of 2  $m/z$ . It appears that the partially reduced products are *not* intermediates to the fully reduced product **2e'**, as an increase in catalyst only results in the loss of **2e** (likely due to hydrosilylation) without an increase in **2e'**. As such, it appears that different pathways generate both reduction products simultaneously. Of the compounds tested, only **1e** exhibited this behavior. The production of multiple reduction products with different numbers of double bonds from the same reactant ultimately prevents assignment of unsaturation state to a reactant of unknown composition.

In summary, it was found that all oxygen-containing groups, with the exception of phenols, are reduced, that aromatic double bonds are stable but that aliphatic double bonds are only preserved in ring systems, and that carbon-carbon bonds are not made or broken. The more complete reduction compared to previous work<sup>12,13</sup> is attributed to the elevated catalyst concentration that is employed in the current method. Excess catalyst has been consistently used across all standards and samples.

**Comprehensive GC tmemGC-TOF-MS.** The distribution of hydrocarbon standards and reduced model compounds in GC×GC space (Figure 1) mapped a chemical landscape that

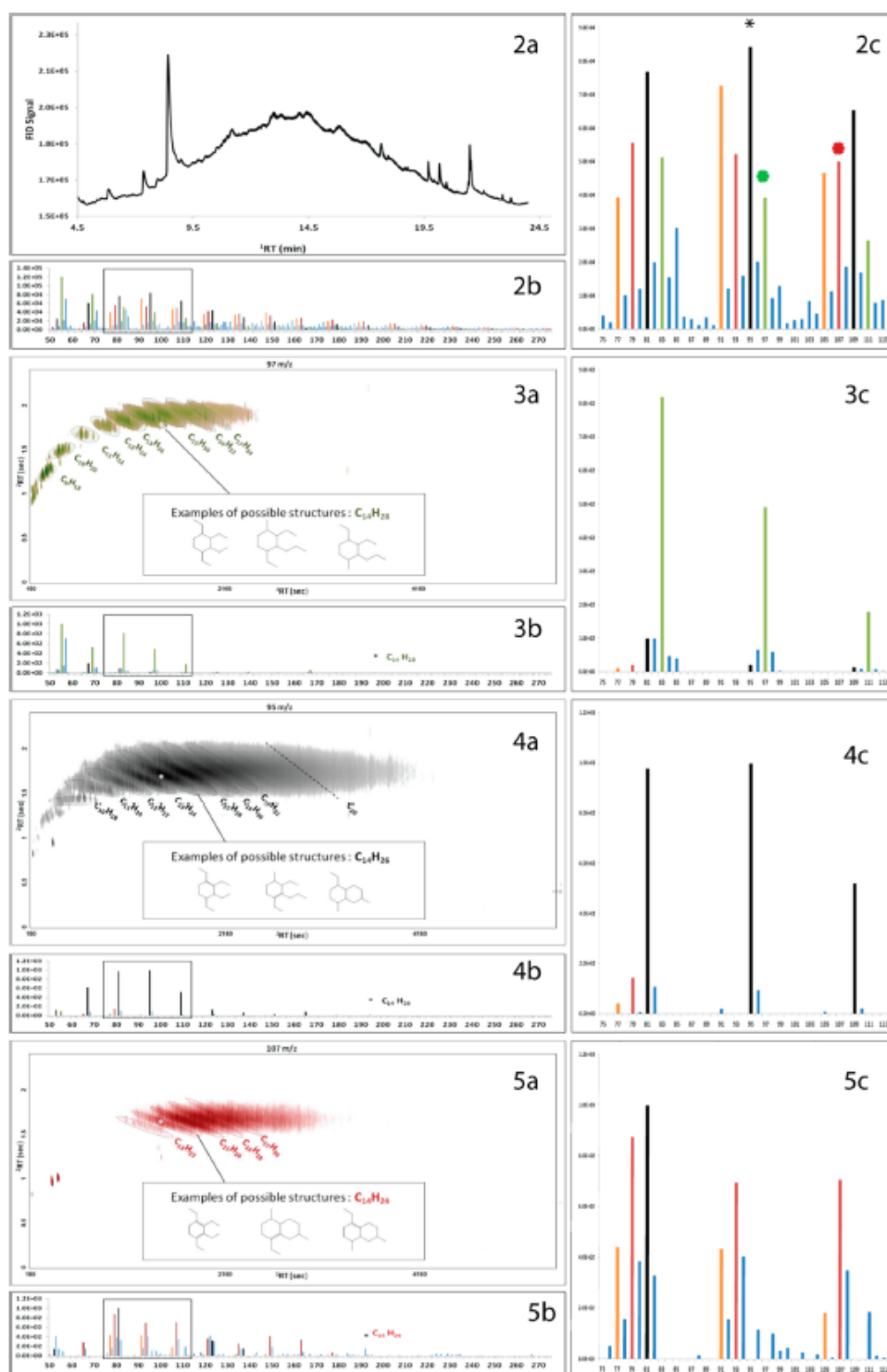


Figure 2–5. GC-FID chromatogram (2a) of reduced SRFA. GC-MS mass spectrum (2b) of material eluting from 10.0–17.6 minutes. Inset in 2b shows (2c) clusters of prominent fragmentation ions. Ions offset by 14  $m/z$  are colored to indicate potential connectivity. Selected ion chromatograms of 97  $m/z$  (3a), 95  $m/z$  (4a) and 107  $m/z$  (5a) display three distinct series within a single GCxGC-TOF-MS run of reduced SRFA.

Figure 2–5. continued

Each series is comprised of alicyclic hydrocarbons, with circled regions indicating isomers of the presented molecular formula. Insets show possible structures for regions. Mass spectra (3b, 4b, 5b) of compounds representative of the series are presented, and insets (3c, 4c, 5c) display prominent fragmentation ions. \*denotes peak location in chromatograms, molecular ion in spectra, prominent ions in 2c

enables preliminary assignment of broad functionality to unknown hydrocarbons based on first and second dimension retention times (<sup>1</sup>RT and <sup>2</sup>RT respectively). Colored areas indicate where saturated (alkane), unsaturated, and aromatic hydrocarbons are expected to elute. A thorough description of the assignment of these regions is provided as Supporting Information.

**Terpenoids in SRFA.** As previously mentioned, we were unable to separate the UCM into discrete peaks using GC-FID/MS. However, GC-MS data still proved to be useful for structural characterization. In a time averaged, baseline subtracted spectrum (Figure 2b) taken along the UCM we observe clusters of prominent fragmentation ions which are spaced 14 *m/z* apart (Figure 2c). Based on previous knowledge of EI-MS hydrocarbon fragmentation we suspected that many closely related compounds contributed these fragments and used this knowledge to further interrogate the sample by GC×GC-TOF-MS.

Three series of compounds are largely responsible for the diversity of fragment ions observed in reduced SRFA by GC×GC-TOF-MS (Figure, 2b and c). Each series contains many individual compounds which occupy a two-dimensional GC×GC space. These three series are displayed in Figures 3a, 4a, and 5a in 97 *m/z*, 95 *m/z*, and 107 *m/z* single ion chromatograms. Within each series individual compounds yield very similar mass spectra, as certain fragment ions are highly conserved across the series. For example, in the hydrocarbon series identified with 97 *m/z* (Figure 3a) we observe three prominent fragmentation ions: 83 *m/z*, 97 *m/z* and 111 *m/z* (Figure 3c). These ions are spaced at 14 *m/z*, the same spacing observed in GC-MS of the UCM, and are present in almost all observed peaks from this series. Compounds identified in Figure 4a have prominent fragment ions of 81 *m/z*, 95 *m/z* and 109 *m/z* (Figure 4c), and compounds in Figure 5a have ions of 79 *m/z*, 93 *m/z* and 107 *m/z* (Figure 5c). All together, these three series represent the prominent fragment ions observed in GC-MS analysis. Additionally, the large molecular size range (<sup>1</sup>RT range) and abundance distribution of each series are similar to the features observed in the GC-FID chromatogram (Figure 2a), and confirms the GC×GC observation that the UCM is largely represented by the three series of compounds presented.

It was further determined that the three series in reduced SRFA are comprised of alicyclic hydrocarbons. Many of the identified peaks have MS with detectable molecular ions, and none of the spectra from any of the series are consistent with hydrosilylated compounds. Model compound work clearly demonstrated that all oxygen containing groups besides phenols are reduced to hydrocarbons. Thus, chemical formulas for these compounds were easily calculated. In the series depicted in Figure 3a, the peaks enclosed by a particular region all have the same chemical formula. From these chemical formulas, it was determined that all three of the series are comprised of hydrocarbons with varying degrees of unsaturation (*n*, where *n* = number of rings plus double bonds). This assignment is also supported by elution of the three series in the same unsaturated hydrocarbon region established in Figure

1. In model compound work it was established that double bonds are only preserved on alicyclic compounds. Therefore, under these conditions, reduction products that are unsaturated hydrocarbons are most likely alicyclic (with or without double bonds).

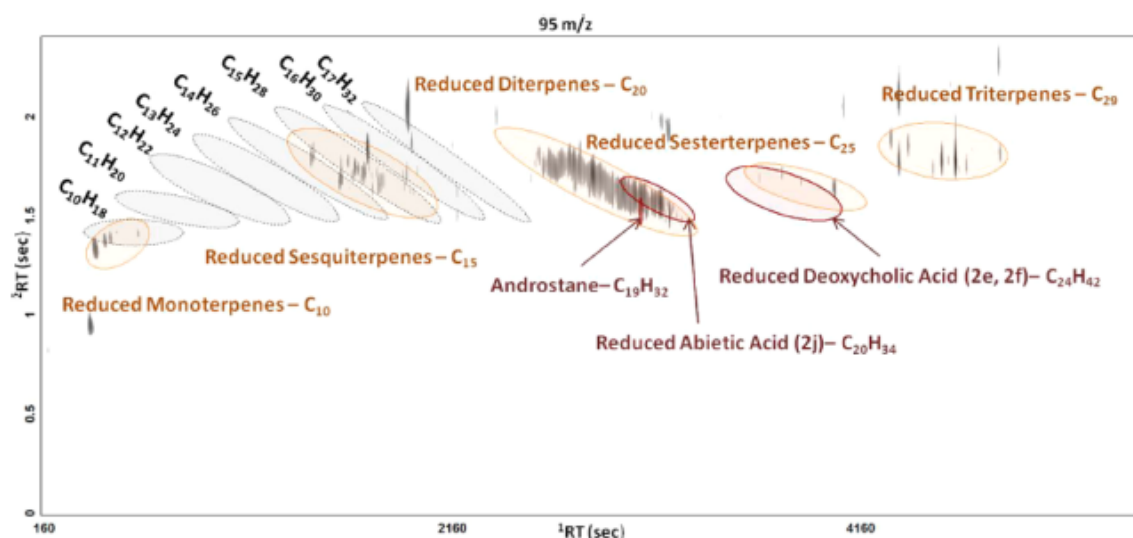
These alicyclic reduction products are notable for a number of reasons. First, they are broadly distributed over a wide molecular size range. The most abundant compounds are centered between C<sub>13</sub> to C<sub>14</sub>, but some material appears to be larger than C<sub>20</sub> (Figure 4a). Unfortunately, the analytical method did not preserve molecular ions for compounds >C<sub>17</sub>. The size-range of compounds observed in the reduced sample has important implications for source material structure, in that terpenes and terpenoids are the only biomolecules with alicyclic backbones of this size. Terpenoids are biomolecules synthesized from isoprene units that contain five carbons. Thus, terpenoid carbon backbones often contain 10, 15, 20, 25, and 30 carbons (although, larger terpenoids can be decarboxylated). By convention, terpenes are strictly hydrocarbons, while terpenoids are oxygenated.<sup>18</sup>

Within a given series, alicyclic reduction products are offset by a CH<sub>2</sub> group (14 *m/z*). This methyl offset is not an artifact of the reduction because saturated C–C bonds are stable to the reduction, and must be a feature of native, unreduced SRFA. In fact, offsets of carboxylic acid groups between individual compounds in Suwannee River Fulvic Acid have been observed using FTICR-MS.<sup>7</sup> When reduced, these offsets would appear as a difference of 14 *m/z*, or the equivalent of a CH<sub>2</sub> group in the molecular formula.

All compounds within a series are unsaturated to the same degree, and each series varies by one degree of unsaturation, or 2 *m/z*. All compounds in Figure 3a have one degree of unsaturation, meaning that they contain one alicyclic ring. Compounds displayed in Figure 4a have two degrees of unsaturation. These compounds may contain a double bond on an alicyclic structure, a fused cyclic structure, or two nonfused rings. Finally, compounds in Figure 5a have three degrees of unsaturation. Possible structures that satisfy each chemical formula and the requirements outlined above are depicted in insets within Figures 3a, 4a, and 5a. Unfortunately, exact structures are unavailable because the product MS are too ambiguous. The recovery of compounds offset by 2 *m/z* (Figures 2b, 3b, and 4b) would corroborate other studies which have identified this feature in SRFA.<sup>7</sup> However, as established in model compound work, the unsaturation state of products is not necessarily reflective of the unsaturation state of reactants. Thus, while the appearance of series offset by degrees of unsaturation is congruent with other previous studies of SRFA, the potential that reduction artifacts contribute to this feature in the current data set must be considered a caveat.

There are a large number of peaks within an enclosed region that share the same molecular ion. Mass spectra of these compounds contain the same fragment ions, but the relative abundances of these ions vary. This confirms that each deconvoluted mass spectrum belongs to a unique compound, but that all compounds within a region are structurally related.





**Figure 6.** 95  $m/z$  selected-ion chromatogram of reduced bald-cypress leaf litter. Dark markings indicate compound peaks. Regions corresponding to reduced terpenes/terpenoids are colored in orange and labeled with size class. Also overlaid are series from Figure 4a (black) and selected reduced compounds and hydrocarbon standard from Figure 1 (red).

Therefore, peaks within a region are very likely to be stereoisomers and/or positional isomers.

Based on this body of evidence, that compounds in reduced SRFA elute within the chromatographic region occupied by unsaturated, alicyclic standard compounds, that double bonds are unlikely to be stable in this reduction system unless they are present within a ring system, that contiguous (no ether or ester bonds) carbon skeletons between  $C_{10}$  and  $C_{17}$  dominate the chromatogram and are not created during the reduction, it can be concluded that the majority of SRFA carbon that is detectable by GC following this reduction (i.e., 12% of the SRFA carbon) resembles cyclic terpenes. This observation is consistent with previous NMR-based studies that have suggested that terpenoids must be present in SRFA<sup>9</sup> but these compounds have previously not been chromatographically isolated. To provide some context to the reduction products observed in SRFA, a potential source material was examined.

**Terpenoids in Bald Cypress Leaves.** In the Suwannee River, a major DOM source is organic matter leached from leaf litter.<sup>19</sup> To confirm the presence of terpenoids in source material we ran the chemical reduction on freshly fallen leaf litter gathered from Tifton, GA. The leaves are of the genus *Taxodium*, primarily from the bald-cypress (*Taxodium distichum*). Bald-cypress is present throughout the entire Suwannee River watershed, and is certain to be a major contributor of DOM to SRFA. Although the potential of (leaf colonizing) bacterial and fungal terpenoids must be considered, the intactness of our sample implies that any observed compounds are ultimately plant derived. The reduced leaf standard was chromatographically separated using GC×GC and displayed using the 95  $m/z$  chromatogram (Figure 6).

The chemical reduction of these source terpenoids to hydrocarbons, and their distribution in the unsaturated region of GC×GC space, confirms that compounds detected in reduced SRFA are indeed derived from terpenes. For example, there are at least two isomers of pinene, a type of bicyclic

monoterpene, which are found in the leaves of conifers<sup>20</sup> and which are present in the extracts of the leaf litter. Under the reaction conditions used, pinene would be expected to reduce to a chemical formula of  $C_{10}H_{18}$ . We find that such products are present, and elute near the classification region of reduced SRFA of the same chemical formula. Conifer leaves are also known to contain bicyclic sesquiterpenes such as  $\alpha$ -Cadinene<sup>21</sup> (found in leaf extracts). The reduction product of such bicyclic terpenes would have the chemical formula  $C_{15}H_{28}$ . Again, we find reduced hydrocarbons of that chemical formula elute in the same GC×GC space as reduced SRFA products. This correspondence does not mean that intact terpenes are present in SRFA. Rather, these smaller terpenoids function as additional model compounds by further establishing the domain of reduced alicyclic compounds in GC×GC space. Such a correspondence is especially reassuring for  $C_{15}$  alicyclic compounds, which are prominent in reduced SRFA, but for which we have no purchased standards.

Consistent with previous studies of *Taxodium* wood,<sup>22</sup> the leaf mixture also contained abundant reduction products of the diterpenoid resin acids and triterpenoid phytosterols. It was previously mentioned that the reduction products of 1c are isomers, which elute over a chromatographic region. This region is overlaid in Figure 6. There are at least four (and potentially eight or greater) diterpenoid<sup>20,23</sup> that are naturally produced by conifers. These other compounds would also be expected to reduce to isomers, and this is confirmed by the numerous isomers of diterpenoid reduction products that appear in the chromatogram of the reduced leaf mixture (Figure 6).

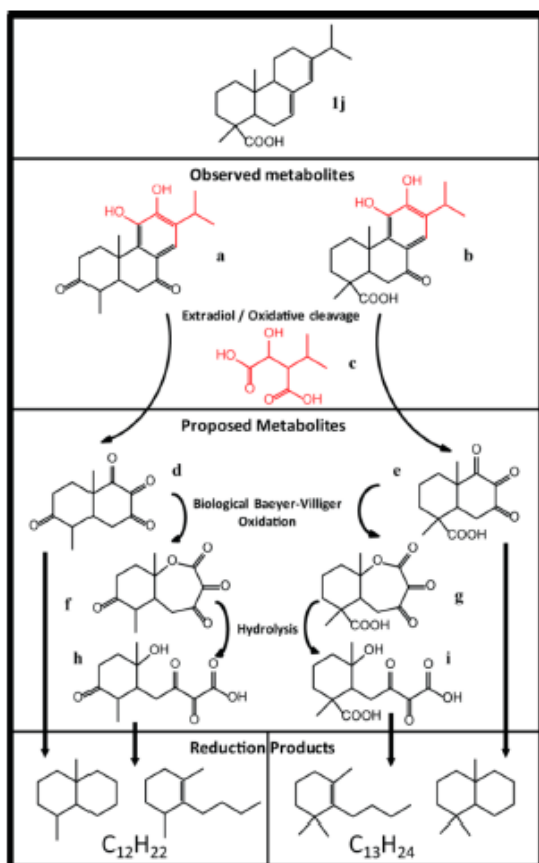
In contrast to the resin acids, the reduction products of triterpenoid phytosterols appear as discrete, but highly abundant, individual peaks. The small number of peaks indicates that phytosterols are less diverse and/or not (as easily) isomerized by reaction conditions. Reduction products of intact resin acids and phytosterols are not identified in the reduction products of SRFA, which suggests that these

compounds are rapidly removed during diagenesis or are transformed into material that escapes detection. Reduced sesterterpenes, which are considered to be rare in nature,<sup>24</sup> are also present in the leaf litter sample. The reduced compounds elute in the same GC×GC space as the reduction products of model compound **1e**.

When considering the possible contribution of terpenoids as source material to SRFA, resin acids are an attractive source, especially considering that the distribution of reduced isomers is similar in shape to that of reduced SRFA. However, there are clearly SRFA reduction products larger than C<sub>20</sub> (Figure 4a), indicating that phytosterols are also likely to be a significant contributor to SRFA. The cyclic portion of larger terpenoids such as carotenoids could also be a source of the alicyclic compounds observed in SRFA. In short, analysis of potential source material confirms that cyclic terpenoids larger than C<sub>15</sub> appear to be a major source for the alicyclic compounds observed in reduced SRFA.

Finally, in Scheme 1 we present a *plausible* pathway by which microbial degradation of parent terpenoid compounds could generate the cyclic hydrocarbons observed in reduced SRFA. We used the resin acid Abietic acid (**1j**) as our precursor model

**Scheme 1. Hypothetical Degradation Pathway from Cyclic Terpenoid to Proposed Metabolites (Components of Native SRFA), Which Then Could Be Further Degraded to Produce the Compounds Observed in Reduced SRFA**



compound. Partially degraded abietanes **a** and **b** have been observed as bacterial metabolites of **1j**.<sup>25</sup> These metabolites only differ by one carbon number, which is the result of decarboxylation at the C-4 position. Further degradation is hypothesized to occur via a c-ring cleavage, catalyzed by a known meta-cleavage dioxygenase,<sup>25</sup> producing 2-isopropyl malic acid (**c**) in the process. However, no degradation products beyond **a** and **b** have been reported on in the literature.

We suggest that further metabolites (**d**, **e**) could exist that are either not amenable to recovery, or are extremely short-lived. Such bicyclic metabolites could degrade further via a biological Baeyer-Villiger oxidation, forming **f** and **g**. The enzyme which catalyzes this reaction is nonspecific and capable of oxidizing a large range of cyclic substrates.<sup>26</sup> These could then be environmentally hydrolyzed, forming **h** and **i**, and/or potentially ester linked into larger compounds. Reduction of these hypothetical metabolites (**d**, **e**, **h**, **i**) could produce carbon backbones of the size and unsaturation observed in reduced SRFA. Although speculative, it appears that degradation via decarboxylation (and then reduction) is the simplest explanation for the observed CH<sub>2</sub> offset of products in reduced SRFA.

**Method Contributions to DOM Analysis.** We have effectively shown that chemical reduction followed by GC×GC analysis of SRFA allows for moderate interrogation of both structural and compositional diversity, where compositional diversity refers to the diversity of molecular formulas within a mixture of compounds, and structural diversity refers to the arrangement of elements within a single compound.

Although capable of revealing compositional diversity in the reaction products, compositional diversity in native SRFA is not directly supported. As the reduction method is comprehensive, it is not possible to determine how many oxygenated functional groups were reduced to form the observed compounds, and therefore it is impossible to extrapolate the exact original composition of reduced compounds. However, we can be certain that direct carbon-carbon bonds were not broken or created during the reduction. The distribution of reduced compounds has a maximum in intensity at a molecular formula of C<sub>14</sub>H<sub>26</sub>. Unreduced SRFA has a peak maximum at around 350 *m/z*, which would correspond to a carbon number per molecule at ~15–20.<sup>27</sup> It is therefore unlikely that individual SRFA compounds would contain more than one degraded terpenoid backbone. If true, then the observed compositional diversity in reduced SRFA is in part a result of the compositional diversity of unreduced SRFA.

There is a great deal of structural diversity present within the UCM, supported by the many isomers identified for each region of a given chemical formula. As a caveat, an unknown amount of diversity could be an artifact of catalyst induced isomerization. However, this effect has only been observed in fused cyclic model compounds. The series displayed in Figure 2a are alicyclic, but contain only one ring and no double bonds. Thus, the many compounds observed within each region are certainly *not* artifacts of the reduction. This demonstrates structural variation within compositionally identical carbon backbones of reduced SRFA, and therefore native SRFA. Broad signals in NMR spectra have been attributed to structural diversity, but to the best of our knowledge this is the first time that structurally distinct isomers have been recovered from SRFA.

The isolation of a significant fraction of terpenoid-derived material indicates that all classes of biomolecules are not preserved equally within humic substances. This observation is striking, considering that terpenoids are likely to account for only a small percentage of total organic matter input. The results presented here indicate that terpenoid derived compounds are preferentially preserved within SRFA. The preferential accumulation of terpenoids in sediments<sup>28</sup> is well-known, but the current findings support long-term accumulation of these biomarkers in the dissolved phase. Further studies are necessary to confirm if this trend is also observed across different ecological settings, including marine environments.

## ■ ASSOCIATED CONTENT

### Supporting Information

Previously published mechanism of chemical reduction; rationale for GC×GC geography; structure rational for produced isomers. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*email: [narakawa@ucsd.edu](mailto:narakawa@ucsd.edu).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank the National Science Foundation for funding (NSF OCE-1155269 to LIA), as well as the following researchers for inspiration, chemical insights, chromatographic support, and foliage: Cameron McIntyre, Vladimir Gevorgyan, Rama Nimmagadda, Brad Allpike, Jack Cochran, Christoph Aeppli, Bob Nelson, Andrew Mehring, and Casey Harris.

## ■ REFERENCES

- (1) *Humic Substances in Soil, Sediment and Water*; Aiken, G. R., Mcknight, D. M., Wershaw, R. L., MacCarthy, P., Eds.; Wiley: New York, 1985.
- (2) Benner, R. Chemical Composition and Reactivity. In *Biogeochemistry of Marine Dissolved Organic Matter*; Hansell, D. A., Carlson, C. A., Eds.; Academic Press: San Diego, 2002; pp 59.
- (3) Leenheer, J. A.; Nanny, M. A.; McIntyre, C. Terpenoids as major precursors of dissolved organic matter in landfill leachates, surface water, and groundwater. *Environ. Sci. Technol.* **2003**, *37* (11), 2323–2331.
- (4) Lam, B.; Baer, A.; Alaei, M.; Lefebvre, B.; Moser, A.; Williams, A.; Simpson, A. J. Major structural components in freshwater dissolved organic matter. *Environ. Sci. Technol.* **2007**, *41* (24), 8240–8247.
- (5) Woods, G. C.; Simpson, M. J.; Simpson, A. J. Oxidized sterols as a significant component of dissolved organic matter: Evidence from 2D HPLC in combination with 2D and 3D NMR spectroscopy. *Water Res.* **2012**, *46*, 3398–3408.
- (6) Kujawinski, E. B.; Hatcher, P. G.; Freitas, M. A. High-resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids: Improvements and comparisons. *Anal. Chem.* **2002**, *74*, 413–419.
- (7) Stenson, A. C.; Marshall, A. G.; Cooper, W. T. Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Anal. Chem.* **2003**, *75*, 1275–1284.
- (8) Ourisson, G. The general role of terpenes and their global significance. *Pure Appl. Chem.* **1990**, *62*, 1401–1404.
- (9) Leenheer, J. A.; Rostad, C. E. *Tannins and Terpenoids As Major Precursors of Suwannee River Fulvic Acid*, U.S. Geological Survey investigations Report 2004-5276; U.S. Geological Survey: Denver, CO, 2004.
- (10) Ertel, J. R.; Hedges, J. I.; Perdue, E. M. Lignin signature of aquatic humic substances. *Science* **1984**, *223* (4635), 485–487.
- (11) Woods, G. C.; Simpson, M. J.; Koerner, P. J.; Napoli, A.; Simpson, A. J. HILIC-NMR: Toward the Identification of individual molecular components in dissolved organic matter. *Environ. Sci. Technol.* **2011**, *45* (9), 3880–3886.
- (12) Nimmagadda, R. D.; McRae, C. R. Characterisation of the backbone structures of several fulvic acids using a novel selective chemical reduction method. *Org. Geochem.* **2007**, *38*, 1061–1072.
- (13) *Advanced Characterization of Natural Organic Matter (NOM) in Australian Water Supplies*, Research report 80; Water Quality Research Australia: Adelaide, SA, 2010.
- (14) Frysinger, G. S.; Caines, R. B.; Xu, L.; Reddy, C. M. Resolving the unresolved complex mixture in petroleum-contaminated sediments. *Environ. Sci. Technol.* **2003**, *37*, 1653–1662.
- (15) Ball, G. I.; Aluwihare, L. I. CuO-oxidized dissolved organic matter (DOM) elucidated by comprehensive two dimensional gas chromatography-time of flight-mass spectrometry (GC×GC-TOF-MS). *Org. Geochem.* **2014**, *87*, 87–98.
- (16) Rubin, M.; Schwier, T.; Gevorgyan, V. Highly efficient  $B(C_6F_5)_3$ -catalyzed hydrosilylation of olefins. *J. Org. Chem.* **2002**, *67*, 1936–1940.
- (17) Peakman, T. M.; Maxwell, J. R. Acid-catalysed rearrangements of steroid alkenes. Part 1. Rearrangement of 5 $\alpha$ -cholest-7-ene. *J. Chem. Soc., Perkin Trans. 1* **1988**, 1065–1070.
- (18) Zwenger, S.; Basu, C. Plant terpenoids: Applications and future potentials. *Biotechnol. Mol. Biol. Rev.* **2008**, 1–7.
- (19) Averett, R. C.; Leenheer, J. A.; McKnight, D. M.; Thorn, K. A. *Humic Substances in the Suwannee River, Georgia: Interactions, Properties, and Proposed Structures*, U.S. Geological Survey Water-Supply Paper 2373; U.S. Government Printing Office: Washington, DC, 1994.
- (20) Gershenzon, G. J.; Dudareva, N. The function of terpene natural products in the natural world. *Nat. Chem. Biol.* **2007**, *3* (7), 408–414.
- (21) Yatagai, M.; Sato, Y.; Takahashi, T. Terpenes of leaf oils from Cupressaceae. *Biochem. Syst. Ecol.* **1985**, *13*, 377–385.
- (22) Zhong, A.; Sleighter, R. L.; Salmon, E.; McKee, G. E.; Hatcher, P. G. Combining advanced NMR techniques with ultrahigh resolution mass spectrometry: A new strategy for molecular scale characterization of macromolecular components of soil and sedimentary organic matter. *Org. Geochem.* **2011**, 903–916.
- (23) Himejima, M.; Hobson, K. R.; Otsuka, T.; Wood, D. L.; Kubo, I. Antimicrobial terpenes from oleoresin of ponderosa pine tree *Pinus ponderosa*: A defense mechanism against microbial invasion. *J. Chem. Ecol.* **1992**, *18*, 1809–1818.
- (24) Cordell, G. A. The occurrence, structure and elucidation and biosynthesis of the sesterterpenes. *Phytochemistry* **1974**, *13*, 2343–2364.
- (25) Martin, V. J. J.; Yu, Z.; Mohn, W. W. Recent advances in understanding resin acid biodegradation: Microbial diversity and metabolism. *Arch. Microbiol.* **1999**, *172*, 131–138.
- (26) Mihovilovic, M. D.; Müller, B.; Stanetty, P. Monooxygenase-mediated Baeyer-Villiger Oxidations. *Eur. J. Org. Chem.* **2002**, 3711–3730.
- (27) Witt, M.; Fuchser, J.; Koch, B. P. Fragmentation studies of fulvic acids using collision induced dissociation fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2009**, 2688–2694.
- (28) Brassell, S. C.; Eglinton, G.; Maxwell, J. R. The geochemistry of terpenoids and steroids. *Biochem. Soc. Trans.* **1983**, *5*, 575–586.

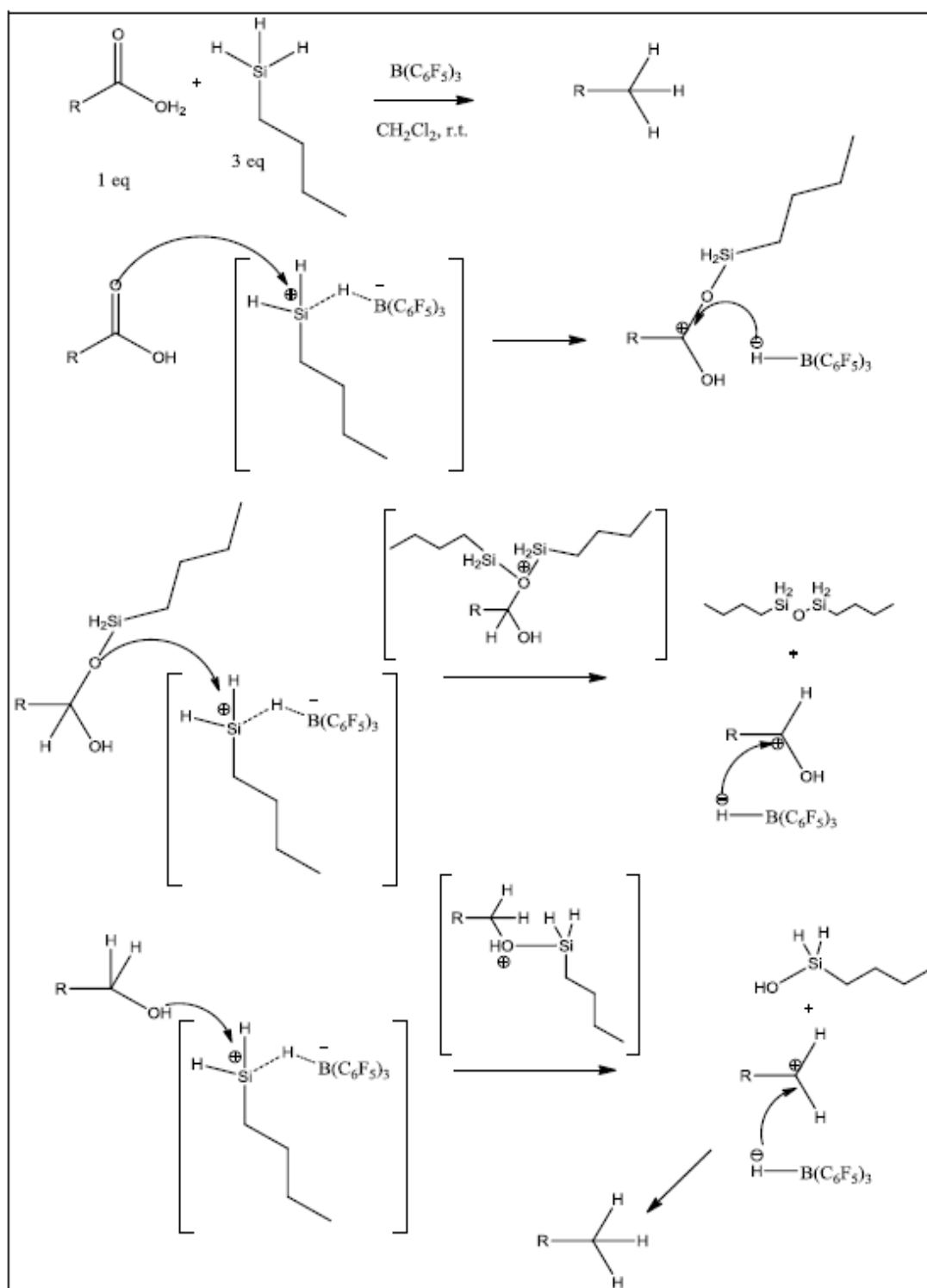


Figure S1. Chemical reduction mechanism as published in Nimmagadda and McRae (1)

**Comprehensive GCxGC-TOF-MS.** The distribution of hydrocarbon standards and reduced model compounds in GCxGC space (Figure 2, main text) mapped a chemical landscape that enabled preliminary assignment of broad functionality to an unknown compound based on its 1<sup>st</sup> and 2<sup>nd</sup> dimension retention time (<sup>1</sup>RT and <sup>2</sup>RT respectively). Separation in the 2<sup>nd</sup> dimension is very strongly driven by chemical affinity to the non-polar phase, such that the least polar compounds elute with later <sup>2</sup>RTs. Of the standards analyzed in this study, the n-alkane series grouped at the top of the chromatogram, effectively delineating the boundary between saturated and unsaturated hydrocarbons. Accordingly, aromatic hydrocarbons have small <sup>2</sup>RTs, and thus group at the bottom of the chromatogram. A wide range of C<sub>12</sub>H<sub>18</sub> (aromatic) isomers with different alkane substituents (boxed compounds in Figure 2) were analyzed to further refine the GCxGC landscape. It is assumed that there is no aromatic isomer of C<sub>12</sub>H<sub>18</sub> that would produce a significantly longer <sup>2</sup>RT than observed in Figure 2, allowing the boundary between aliphatic and aromatic compounds in GCxGC space to be established. An important advantage of GC techniques is that many stereoisomers (cis and trans decahydronaphthalene, for example) and positional isomers can be separated without the use of chiral columns. In summary, the <sup>1</sup>RT and <sup>2</sup>RT spatial distribution of standards shown in Figure 1 lays the foundation for enabling robust assignment of detected peaks to compound classes, and complements and confirms structural assignments based on TOF mass spectra when NIST EI-MS data are not available.

#### **Validation of proposed structures as reduction products of Deoxycholic Acid:**

It has been observed that multiple products are formed during the reduction of Deoxycholic Acid 1j. The following experiment tested the effect of varying catalyst concentration on the products

that are formed during the reduction. The exact structure of products was not determined, as it was not critical to the overall analysis.

**Experimental:** Deoxycholic Acid **1j** was reduced as outlined in the main text under materials and methods. Specific substrate and catalyst amounts used were as follows: **Low Lewis Acid Catalyst:** 8 mg **1j**, 8 mg. Tris(pentafluorophenyl)borane, 100uL nBS. **High Lewis Acid Catalyst:** 8 mg **1j**, 16 mg. Tris(pentafluorophenyl)borane, 100uL nBS. Work-up for NMR and GC-MS/FID followed experimental protocols described in the main text. For NMR, samples were dried under N<sub>2</sub> and resuspended in CDCl<sub>3</sub> before analysis

NMR experiments were performed on a Varian Inova™ 500 MHz spectrometer with a 5 mm inverse detect probe. For <sup>1</sup>H 1D experiments, 64 scans were acquired with 30k time domain points. For <sup>13</sup>C 1D experiments, 100000 scans were acquired with 58K time domain points and a recycle delay of 2 s. For heteronuclear single quantum coherence (HSQC) experiment, 80 scans were acquired with 2k data points in the F2 dimension and 128 increments in the F1. The <sup>1</sup>J <sup>1</sup>H-<sup>13</sup>C value was set to 140 Hz and the relaxation delay was 2.5 s.

#### **Results/Discussion:**

1. Observed peaks represent actual isomers (as opposed to chromatographic artifacts) produced during reaction

In Figure S.2 both high and low catalyst reduction conditions produce varying amounts of peaks 1 and 2, and similar amounts of peak 3(as quantified with GC-FID). Electron

impact mass spectra of these peaks (Fig S.3) indicate that peaks 1 and 2 represent unsaturated versions of the expected product, while peak 3 represents a fully saturated product (based on molecular ion). Unreduced **1j** contains 2 hydroxyl groups, located at C-3 and C-12. Peaks 1 and 2 result from the elimination of one of these hydroxyl groups, producing a double bond. To verify that peaks 1 and 2 are in fact isomers,  $^1\text{H}$  NMR spectra were collected on both reduction mixtures. The overlaid spectra (Fig. S.4) are very similar, but show different relative intensities of peaks, which is expected since the two mixtures contain different amounts of peaks 1 and 2.

2. In each isomer the double bond is located between two tertiary carbons (i.e. there are no olefinic protons)

When either the C-3 or C-12 hydroxyl group is eliminated the double bond located at either site will contain at least one olefinic proton. However,  $^1\text{H}$  NMR spectra show resonances at 2.7 and 2.9 ppm, which are more upfield than would be expected for an olefinic proton. However, we did not rely on this fact alone to rule out this possibility. The  $^{13}\text{C}$  NMR (Fig S.5 f1 axis) spectrum of the Low Catalyst reduction shows 4 peaks between 120-150 ppm, which are within the range of  $\text{sp}^2$  hybridized carbons. Heteronuclear single quantum coherence (HSQC) spectroscopy (Fig S.5) shows direct H-C coupling indicated as either blue (CH and  $\text{CH}_3$ ) or red ( $\text{CH}_2$ ). The two previously mentioned  $^1\text{H}$  NMR resonances (2.65, 2.9 ppm) correlate to  $\text{sp}^3$  hybridized carbons further upfield, and the 4 downfield carbons (124, 136, 138.5, 139.5 ppm) do not show any correlations to  $^1\text{H}$  NMR resonances. These 4 downfield carbon resonances confirm that peaks 1 and 2 each contain one double bond, and that each is positioned between two tertiary carbons (total of 4 peaks).

3. Lewis Acid catalyst supports double-bond migration as well as methyl shifts, resulting in backbone rearrangement.

The presence of double bonds that are not at the position of OH elimination following the reduction suggests that double bonds are migrating in the ring system. It has been shown that  $\text{BF}_3\text{-OEt}$ , a Lewis acid of similar strength to the  $\text{B}(\text{C}_6\text{F}_5)_3$  catalyst used in the current reaction, catalyzes double bond migration as well as methyl shifts in similar steroid systems<sup>2</sup>. The mass fragmentation patterns helped to further limit where the location of the double bond. Published fragmentation patterns of similar steroids indicate that a double bond located on either the a, b, or c ring of a steroid structure will yield a characteristic 215 m/z fragment ion<sup>3,4</sup>. This fragment ion was not abundant in the spectra of either Peak 1 or 2, suggesting that the double bond is on the d ring. Similar unsaturated steroids with the double bond on the d ring (between either C-13 and C-17, C-14 and C-15, or C-17 and C-20) have 257 m/z as the base peak<sup>2,5</sup>. Peaks 1 and 2 both have 257 m/z as their base peak, and have no olefinic protons. Structures provided in Fig. S.2 satisfy both GC-MS and NMR constraints.



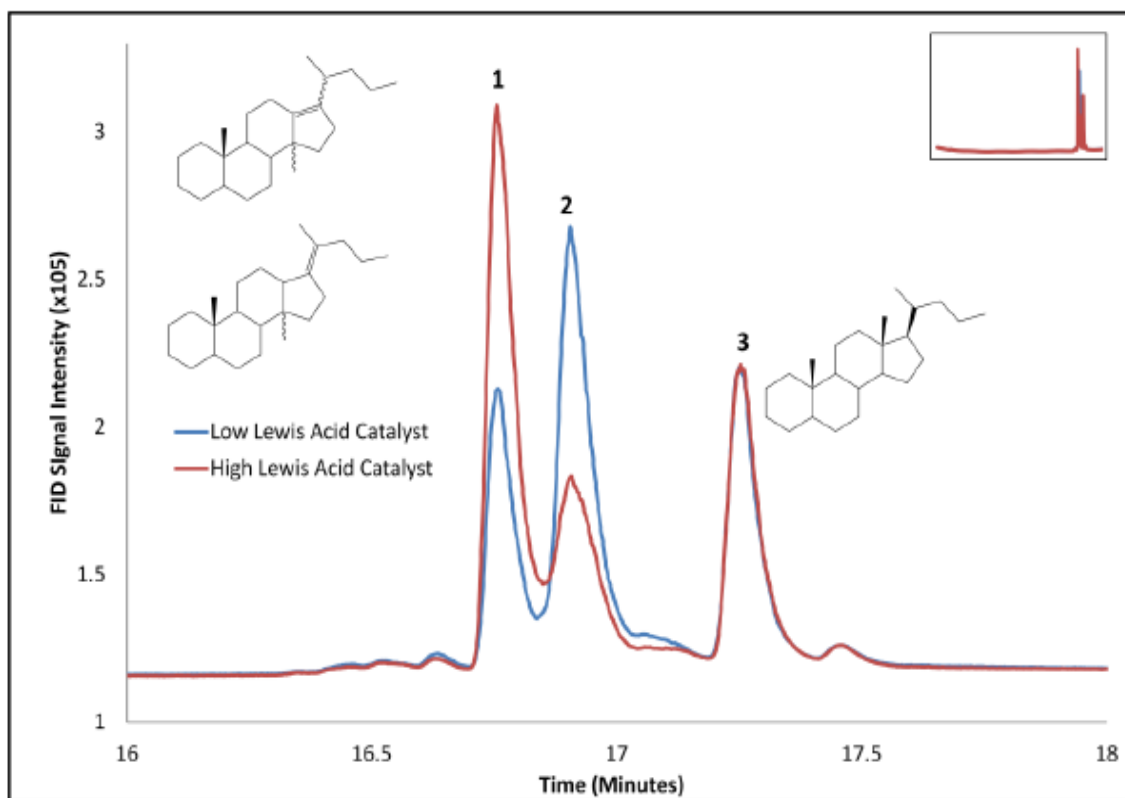


Figure S2. Overlaid GC-FID chromatograms of concurrent reductions with varying catalyst concentrations. Possible structures for peaks 1 and 2 are depicted on left. Structure of peak 3 is depicted on right.

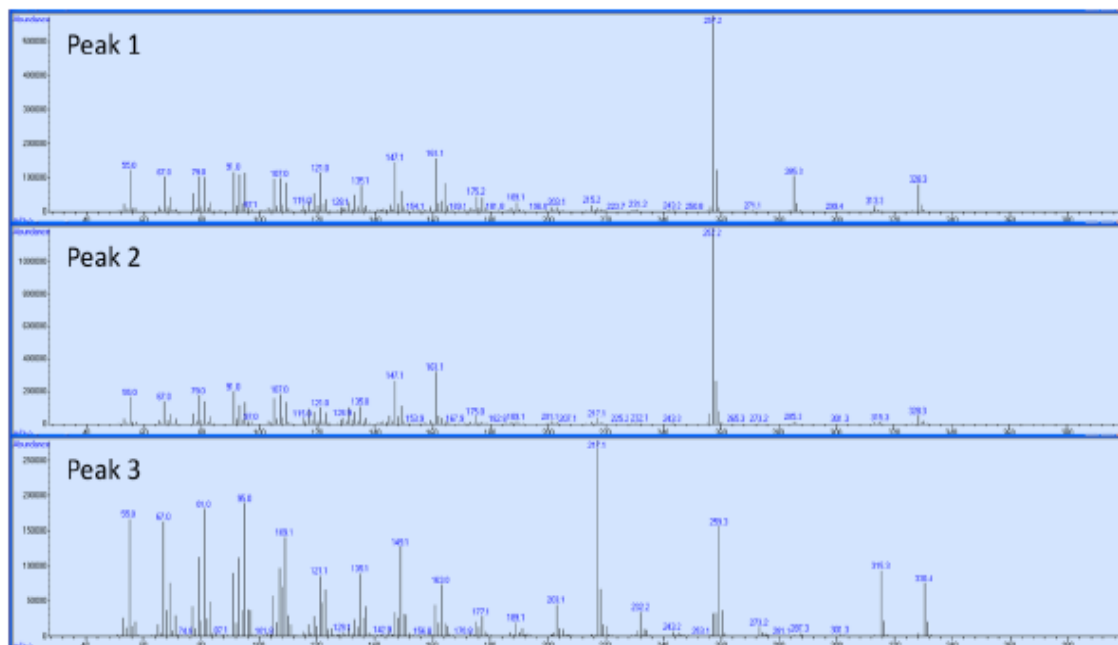


Figure S.3 GC-MS spectra of peaks 1, 2, and 3 from Fig. S.2.

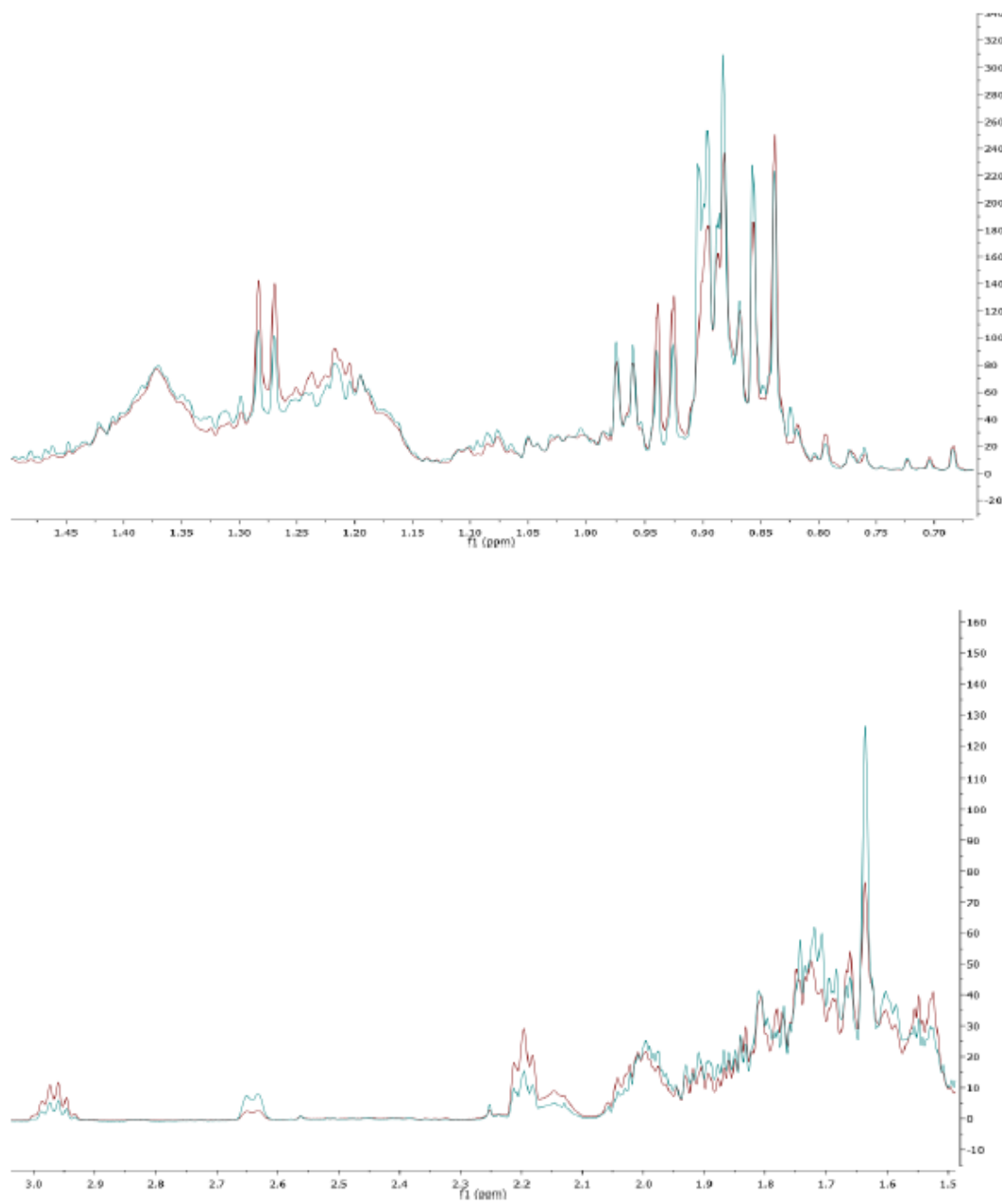


Figure S.4  $^1\text{H}$ NMR of low catalyst reduction mixture (blue) and high catalyst reduction mixture (red). Varying abundances of peaks 1 and 2 (Fig. S.2) are expressed as intensity differences in peaks of the same chemical shift.

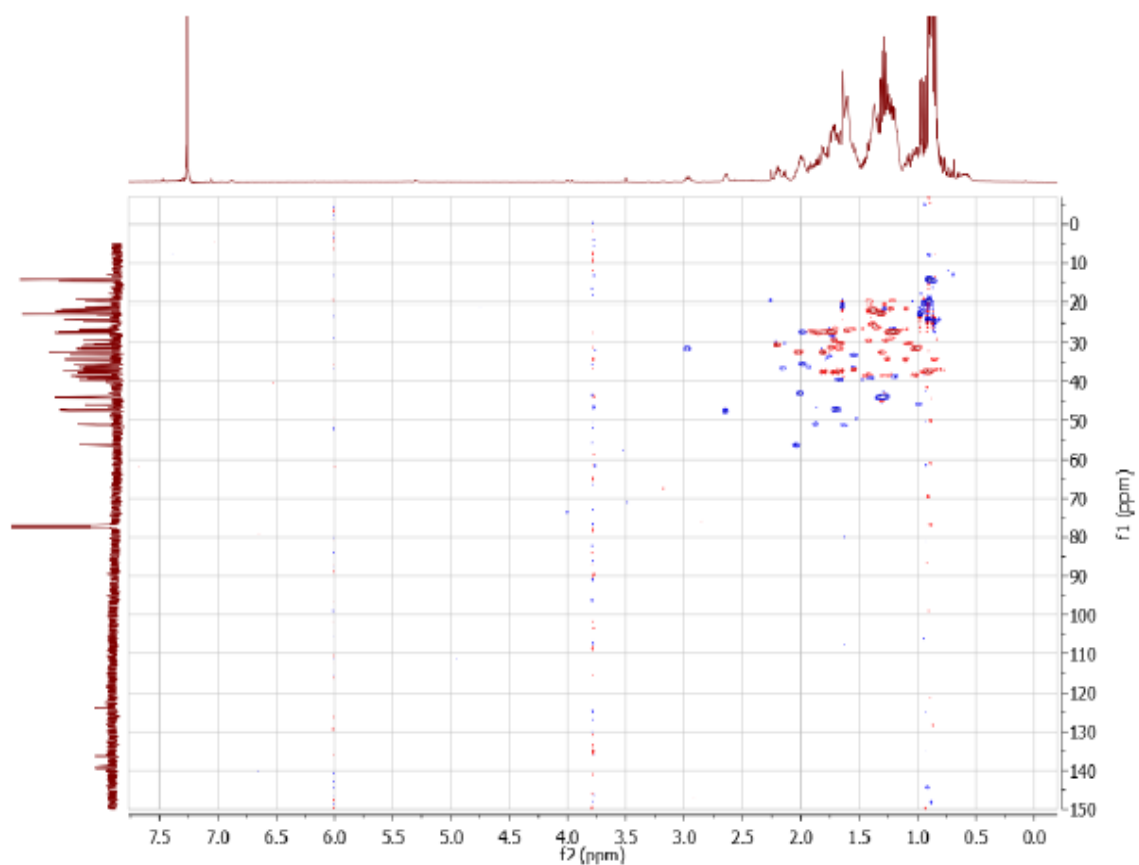


Figure S.5. Heteronuclear Single Quantum Coherence Spectroscopy (HSQC) spectrum of low level catalyst reduction mixture.  $^1\text{H}$ NMR (Fig S.4) is located on the f2 axis.  $^{13}\text{C}$ NMR is located on the f1 axis. Correlations are depicted as either blue (CH, CH<sub>3</sub>) or red (CH<sub>2</sub>). Note that the 4 downfield carbon peaks (120-140 ppm) do not show correlations with  $^1\text{H}$ NMR peaks. The two potential olefinic protons (2.7 and 2.9 ppm), correlate to upfield  $\text{sp}^3$  hybridized carbons.

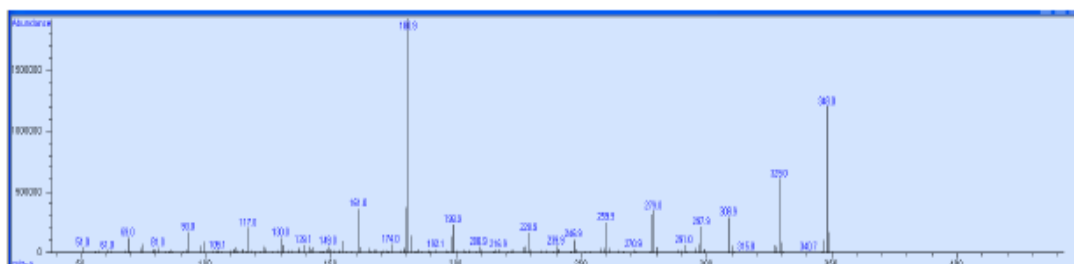


Figure S.6 GC-MS spectrum of catalyst degradation peak

#### Literature Cited

1. Nimmagadda, R.D., McRae, C., 2006. A novel reduction of polycarboxylic acids into their corresponding alkanes using n-butylsilane or diethylsilane as the reducing agent. *Tetrahedron Letters* 47, 3505-3508
2. Peakman, T.M; Maxwell, J.R. Acid-catalysed Rearrangements of Steroid Alkenes. Part 1. Rearrangement of 5 $\alpha$ -Cholest-7-ene. *J. Chem. Soc. Perkin Trans. 1.* **1988**, 1065-1070
3. Partridge, L.G.; Midgley, I.; Djerassi, C. Mass Spectrometry in Structural and Stereochemical Problems. 249. Elucidation of the course of the Characteristic Ring D Fragmentation of Unsaturated Steroids. *J. Am. Chem. Soc.*, **1977**, 7686-7695
4. Zaretskii, V.I.; Vul'fson, N.S.; Zaikin, V.G.; Papernaya, I.B. Mass Spectrometry of Steroid Systems. *Chemistry of Natural Compounds* **1967** 320-327
5. Rubenstein, I.; Sieskind, O.; Albrecht, P. Rearranged Sterenes in a Shale: Occurrence and Simulated formation. *J. Chem. Soc. Perkin Trans. 1.* **1975**, 1833-1836

**Acknowledgments**

Chapter II, in full, is a reprint of the material as it appears in Direct identification of diverse alicyclic terpenoids in Suwannee River Fulvic Acid in *Environmental Science and Technology*, 2015, Arakawa, Neal; Aluiwhare, Lihini, 49: 4097-4105. The dissertation author was the primary investigator and author of this paper.

## **Chapter III**

Carotenoids are the likely precursor of a major fraction of dissolved organic matter

## **Abstract**

Dissolved organic matter (DOM) that can persist for millennia represents an essential pathway of the Ocean's Biological Carbon Pump. The origin of DOM's persistence has never been fully explained, and so, its sensitivity to climate change remains unconstrained. Here, we show, using two independent analytical approaches, that accumulating DOM contains aliphatic compounds with cyclic head groups that are consistent with highly oxidized carotenoid structures. This comprehensively characterized fraction of DOM had a  $^{14}\text{C}$  age of  $\sim 1625$  years, and carotenoid derivatives comprise 10% to 50% of its carbon. The reported chemical modification of carotenoids through oxidation implies that intrinsic stability dictates the persistence of this DOM fraction in the ocean.

## **Introduction**

The Ocean contains approximately 700 Gt of carbon (Siegenthaler & Sarmiento 1993) in the form of dissolved organic matter (DOM), a reservoir that has been implicated in modulating Earth's past climate (Sexton et al. 2011). A significant fraction of this reservoir is comprised of a radiocarbon ( $^{14}\text{C}$ )-depleted component that exists mixed throughout the entire water column (Williams & Druffel 1987). How this component functions as a capacitor for carbon dioxide on different timescales remains unknown because its composition is still only poorly defined. A variety of methods can isolate  $^{14}\text{C}$ -depleted DOM from seawater (Repeta & Aluishiare 2006, Dittmar et al. 2008, Flerus et al. 2012, Lechtenfeld et al. 2014), but existing analytical techniques have thus far struggled to provide definitive chemical structures. Such information is necessary to



determine the source, ultimate fate and reason for DOM accumulation in the ocean on millennial timescales.

Over the last two decades a consensus has emerged identifying terpenoids as major components of terrestrial DOM (Leenheer et al. 2003, Hertkorn et al. 2006, Lam et al. 2007, Arakawa & Aluwihare 2015), and recently molecular level analytical techniques provided unambiguous structural support for their prevalence (Arakawa & Aluwihare 2015). Additionally, comparison of surface and deep marine samples has led investigators to conclude that this class of compounds constitutes part of refractory DOM. Carotenoids, a narrow class of terpenoids, play numerous roles in marine and terrestrial environments including as light-harvesting and photoprotective pigments. Of the 100s of unique marine carotenoids, many share two defining features: alicyclic headgroups of a specific carbon arrangement, and conjugated isoprene units (Liaaen-Jensen 1991). Both of these features have been definitively identified in the present study using a combination of nuclear magnetic resonance spectroscopy (NMR) and mass spectrometry (MS). Such a finding confirms that many of the terpenoid features detected in marine DOM arise from compounds resembling carotenoids.

## **Materials and Methods**

### *Materials.*

All solvents (ACS grade or better), model compounds and hydrocarbon standards were purchased commercially, and used as supplied. n-Butylsilane (n-BS) and Tris(pentafluorophenyl)borane ( $B(C_6F_5)_3$ ) were purchased from Sigma Aldrich, and

stored in an Argon filled desiccator when not in use.  $\text{CH}_2\text{Cl}_2$  was stored over activated molecular sieves (3 Å, 10% wv).

#### *DOM Isolation.*

The isolation protocol was adapted from Dittmar et al. (2008). 330L of seawater from the Scripps Institution of Oceanography Pier (UCSD, La Jolla), were collected and pre-filtered through an AcroPak™ 0.8/0.2 µm polyethersufone (PES) Supor® membrane filter. The filtrate was acidified to pH 2 with conc. Hydrochloric acid (HCl), and passed over activated (1 cartridge volume methanol (MeOH)) 1 gm Agilent Bond Elut PPL styrene-divinylbenzne polymer cartridges @ ~15ml/min. 9 cartridges were used and eluted after 20 L. Prior to elution, cartridges were rinsed with 2 cartridge volumes 0.01 M HCl, followed by 2 cartridge volumes pure water (MilliQ). The MilliQ rinse was done to limit methylation of DOM under MeOH elution. The cartridges were then dried under  $\text{N}_2$  gas, and eluted with 2 cartridge volumes of MeOH. The combined eluents were dried under  $\text{N}_2$  gas, resuspended in MilliQ, frozen, and lyophilized. The total weight of recovered DOM was 235 mg.

#### *DOM Fractionation/purification.*

PPL-DOM was hydrolyzed with 2M HCl at 90°C overnight. The dried hydrolysate was re-dissolved in acidic water (0.01 M HCl) and extracted against ethyl acetate to remove highly lipophilic material. The extraction was repeated twice with 0.01 M HCl and the acidic water layers were combined and subsequently freeze dried. The dry, polar PPL-DOM hydrolysate was resuspended in 0.02 M HCl, and extracted again using a 1 g PPL cartridge. The permeate, and successive acid and MilliQ rinses, were collected and

freeze dried. The permeate plus wash fraction contained primarily sugars and amino acids as determined by NMR and GC-MS data (not presented). Finally, the PPL cartridge was dried, and eluted with MeOH and this eluent is identified as the Carotenoid-Rich fraction in the text, tables (3.1) and Figures (3.8).

#### *Radiocarbon and elemental analysis*

Radiocarbon measurements were generated at the W.M Keck Carbon Cycle Accelerator Mass Spectrometry Laboratory at UC Irvine. Sample graphitizing backgrounds have been subtracted based on combusted glycine and glycine. Sample processing backgrounds have been subtracted based on processed glycine and glycine. All results have been corrected for isotopic fractionation according to the conventions of Stuiver and Polach (1977), with  $\delta^{13}\text{C}$  values measured on prepared graphite using the AMS spectrometer. These can differ from  $\delta^{13}\text{C}$  of the original material, if fractionation occurred during sample graphitization or the AMS measurement, and are not shown. Independent elemental and stable isotopic ( $\delta^{13}\text{C}$ ,  $\delta^{15}\text{N}$ ) characterization of samples was performed using standard elemental analyzer isotope ratio mass spectrometry (EA-IRMS) at the Scripps Institution of Oceanography.

#### *Gas Chromatography-Flame Ionization Detection/Mass Spectrometry (GC-MS/FID)*

An Agilent 7890A Gas Chromatograph system coupled simultaneously to an Agilent 5975C quadrupole mass spectrometer and a flame ionization detector was used for GC-MS/FID characterization of the sample. Splitless injection with 1  $\mu\text{L}$  of analyte was used. Separation was performed on a 5% phenyl poly(dimethylsiloxane) column (J&W

123-5731DB-ht, 30 m, 320  $\mu\text{m}$  i.d., 0.1  $\mu\text{m}$  film) with a temperature program from 50  $^{\circ}\text{C}$  (hold time 0.5 min) to 140  $^{\circ}\text{C}$  (at 10  $^{\circ}\text{C}/\text{min}$ , hold time 0 min) to 320  $^{\circ}\text{C}$  (at 15  $^{\circ}\text{C}/\text{min}$ , hold time 10 min). Helium was the carrier gas with a constant flow of 1.8 mL/min. After separation, the effluent was split between the flame ionization detector (operating at 310 $^{\circ}\text{C}$ ) and the mass spectrometer (70eV ionization, scanning 50-750 m/z).

Quantification of reduced model compounds and reduced SIO Pier DOM was accomplished using a calibration curve generated using a hydrocarbon standard (decahydronaphthalene) and assuming a constant FID response factor across unknown hydrocarbons. When available, reduced model compound structures were confirmed by comparison to the NIST Mass Spectral Standard Reference Database. For SIO pier DOM, manual integration was used to remove peaks associated with catalyst degradation.

#### *Comprehensive Two-Dimensional Gas Chromatography (GCxGC-TOF-MS)*

The GCxGC instrument used is a LECO Pegasus 4D GCxGC-Time-of-Flight Mass Spectrometer (GCxGC-TOF-MS). The term GCxGC refers to the use of two distinct columns in series that have different chemical selectivity. Compounds are separated primarily by volatility in the first column and polarity in the second. The analysis is comprehensive because all of the effluent from the first column is cryofocused and transferred onto the second column. Effluent from the second column is then analyzed by the TOF-MS, which benefits from high spectral acquisition speed (50 Hz-500 Hz). Finally, the data is compiled into a two-dimensional chromatogram that is visualized and processed by ChromaTOF<sup>®</sup> software. Both columns are housed within an Agilent 7890A Gas Chromatograph. The splitless inlet temperature is set at 300 $^{\circ}\text{C}$ . The first

dimension column is a semi-polar Crossbond<sup>®</sup> diphenyl dimethyl polysiloxane column (Restek Rxi-17Si, 30m length, 0.2550mm i.d., 0.25 $\mu$ m film thickness). The column was programmed to remain isothermal at 40 $^{\circ}$ C for 1 min, and ramped to 315 $^{\circ}$ C at 3 $^{\circ}$ C per minute. The modulator temperature was offset by +30 $^{\circ}$ C to the primary oven. The secondary oven (within the GC) housed the second dimension non polar Crossbond<sup>®</sup> dimethyl polysiloxane column (Restek Rxi-1, 1.58m length, 0.250mm i.d., 0.25 $\mu$ m film thickness). The secondary oven temperature was offset by +25 $^{\circ}$ C to the primary oven. The modulation period was 2.5 s, with a hot pulse time of 1.05 s and a cool time of 0.2 s. The carrier gas was Helium at a constant flow of 1.5 mL/min. The acquisition delay on the TOF-MS was set to 160s. The acquisition rate was set to 50Hz, with a range of 5-1000 m/z. Electron Ionization was run at 70eV. Again, reduced model compound structures were confirmed by comparison to the NIST Mass Spectral Standard Reference Database.

#### *Reduction Procedure*

The reduction procedure directly follows Arakawa and Aluwihare (2015) which was modeled after Nimagadda et al. (2007). For model compound reductions, compounds (~2.5 mg) were transferred by syringe to a single flame dried 2 mL vial equipped with a stir bar. Under Argon atmosphere, 10 mg of  $B(C_6F_5)_3$  was added (100  $\mu$ L of 100 mg  $B(C_6F_5)_3$ /1 mL dichloromethane solution). Immediately after addition, 100  $\mu$ L of n-BS was added, also under Argon. The mixture was then allowed to stir-overnight. Note that the sample is completely soluble in dichloromethane after the reduction. Following reduction the samples were treated to remove excess catalyst and siloxane by-products. Approximately 50  $\mu$ L of the reaction mixture was transferred into 250  $\mu$ L of conc.

KOH/MeOH. After 1 minute the mixture was extracted with pentane (3x 100 $\mu$ L). The pentane fractions were collected, washed with H<sub>2</sub>O, and dried over Na<sub>2</sub>SO<sub>4</sub>. The organic fraction was then dried under N<sub>2</sub> to 100 $\mu$ L for GC-FID/MS and GCxGC-TOF-MS analysis. This complete method was reproduced during the reduction of SIO Pier PPL-DOM, with 5 mg. of starting material. See Arakawa and Aluwihare (2015) for additional discussion of the reduction method.

### *NMR Analysis*

Sample (50 mg) was re-suspended in 1 mL of deuterated methanol (CD<sub>3</sub>OD). Samples were analyzed using a Bruker Avance III HD 500 MHz NMR spectrometer equipped with a <sup>1</sup>H-<sup>13</sup>C-<sup>15</sup>N 5 mm, triple resonance inverse (TXI) cryoprobe.

Heteronuclear Single Quantum Coherence (HSQC) spectra were collected in phase-sensitive mode using Echo/Anti-echo gradient selection, sensitivity enhancement and multiplicity editing during the selection step. All 180° carbon pulses were performed using matched adiabatic pulses for inversion and refocusing. 512 scans were collected for each of the 128 increments in the F1 dimension. 1024 data points were collected in F2, a <sup>1</sup>J <sup>1</sup>H-<sup>13</sup>C value of 145. The F2 dimension was multiplied by an exponential function corresponding to a 15 Hz line broadening, while the F1 dimension was processed using a sine-squared function with a  $\pi/2$  phase shift and a zero-filling factor of 2.

Heteronuclear Multiple Bond Correlation (HMBC) were carried out in phase-sensitive mode using Echo/Anti-echo gradient selection (Cicero et al. 2001) and a relaxation optimized delay of 25 ms for the evolution of long-range couplings (Lam et al. 2007, Simpson et al. 2011). 2048 data points were collected in F2 over 1024 scans for

each of the 128 increments in the F1 dimension. The F2 dimension was multiplied by an exponential function corresponding to a 15 Hz line broadening, while the F1 dimension was processed using a sine-squared function with a  $\pi/2$  phase shift and a zero-filling factor of 2.

Spectral predictions were carried out using Advanced Chemistry Development's ACD/NMR Workbook using Neural Network Prediction algorithms (version 2015.2.5). Parameters used for prediction including spectral resolution, and base frequency were chosen to match those of the real datasets as closely as possible. Due to relative fast relaxation in the DOM sample the optimal evolution delay was determined to be 25ms in HMBC experiments. Longer delays theoretically allow longer range couplings to evolve however in DOM fast relaxation leads to the loss of signal if longer evolution delays are used. As such 25ms represents the optimal compromise for the sample that provides the best balance between signal and correlations that can be observed (Lam et al. 2007). However, a 25ms delay corresponds to an  $(1/2J_{\text{H-C}})$  of 20Hz which will bias the stronger and shorter range correlations in the sample, very long range couplings such as 4 and 5 bond correlation which take a long time to evolve will not be detected in DOM. To account for this and permit the simulations to reflect the real situation as closely as possible only  $^1\text{H}-^{13}\text{C}$  couplings greater than 6Hz were included in the calculations. Both 2 bond and 3 bonds correlations were permitted but the two bonds correlations that generally build up the quickest (i.e. will be detected preferentially in DOM due to the short evolution delay that had to be used) were weighted 2:1 over the 3 bond correlations. The result is that the real HMBC data for DOM will bias strong and short correlations

and this is matched, as best possible, by the simulations where strong couplings ( $>6\text{Hz}$ ) and short range (preference of 2 bond over 3 bond) are also given preference.

## Results and Discussion

Marine DOM for this study was isolated by solid phase extraction (SPE) with PPL resin (Dittmar et al. 2008) from surface ocean waters near the Scripps Institution of Oceanography (SIO) and the isolated component represented 39.6% of measurable dissolved organic carbon (DOC). This sample, termed PPL-DOM, has representative compositional and isotopic characteristics (Table Fig. 3.1, Hertkorn et al. 2006). It also represents a temporally consistent fraction of the SIO DOM pool as evidenced by a  $^1\text{H}$ NMR spectrum that is nearly identical to other PPL samples from a year-long, monthly time series at this location (Fig. 3.1). To characterize PPL-DOM at the molecular level a chemical reduction (Nimmagadda & McRae, 2007, Arakawa & Aluwihare 2015) was employed to first make its constituents amenable to separation by gas chromatography (GC). The reduction converts most oxygen containing functional groups (e.g. esters, carboxylic acids, alcohols) to their respective hydrocarbon backbones.

Based on GC-based quantification the chemical reduction of SIO Pier PPL-DOM yielded 10% of its original carbon as GC-amenable products. Calculation is based on quantitative FID analysis using decahydronaphthalene as a hydrocarbon standard, and manual integration of reduction products. Further interrogation of PPL-DOM with comprehensive GCxGC MS (Fig. 3.4) revealed several related series' of isoprenoid-based, alicyclic hydrocarbons very similar to those previously reported in terrestrial DOM (Arakawa & Aluwihare 2015). It was previously reported that for a terrestrial



DOM sample, the most abundant reduction products were attributed to three different series of alicyclic hydrocarbons (Arakawa & Aluwihare 2015). To arrive at this conclusion a systematic evaluation was conducted with model compounds containing different types of functional groups. These compounds were reduced using the same protocol as described here and the number, composition, and yield of the product(s) of these reactions were then used in combination to interpret the outcome of the PPL-DOM reduction. First, during the reduction most organic compounds are reduced completely to hydrocarbons. Second, the chemical formula (based on molecular ions,  $M^+$ ) of reduced hydrocarbons indicated unsaturation in the form of rings or double bonds. Third, double bonds were found to be stable to the reduction only in alicyclic (ring) systems. In the reduced terrestrial DOM, as with marine PPL-DOM, each observed series contained a multitude of compounds with the same degree of unsaturation (equal to sum of rings and double bonds), but varied in overall size. In both types of samples the most prominent series contained hydrocarbons with two degrees of unsaturation, which were attributed to either hydrocarbons with an alicyclic ring and double bond, or bicyclic hydrocarbons. Exact chemical structures could not be supported by mass spectra alone.

The determination of reduction products observed by MS as specifically alicyclic is based primarily on the observed chemical formulae of reduction products and knowledge of the reaction mechanism revealed from reducing standard compounds. As with terrestrial DOM (Arakawa & Aluwihare 2015), the reduction of marine DOM uncovered remarkable structural homogeneity and large isomeric diversity consistent with inferences made from Fourier Transform Ion Cyclotron MS studies (Stenson et al. 2003, Witt et al. 2009, Dittmar 2015).

Two prominent reduction products (Fig. 3.4, peaks A and B) were structurally characterized by matching their retention times ( $^1\text{RT}$ ,  $^2\text{RT}$ ) and fragmentation patterns with those of reduced model compounds (Fig. 3.2). Their carbon skeletons are entirely consistent with terpenoids, and specifically carotenoids. For example, known carotenoid degradation products loliolide and abscisic acid (Taylor & Burden 1970) would produce similar compounds when reduced (Fig. 3.3). The detection of carotenoid-like head group structures in the chromatogram (Fig. 3.4) unambiguously confirms previous NMR reports implicating oxidized terpenoids (Hertkorn et al. 2006) and specifically compounds resembling carotenoids (Lam et al. 2007) as major components of DOM.

Characteristics of the remaining hundreds of ions in reduced PPL-DOM (Fig. 3.4) were also scrutinized to identify further support for the presence of carotenoid degradation products. Hypothetically breaking carbon-carbon bonds in a generic carotenoid (Fig. 3.4) can result in unsaturated alicyclic hydrocarbons between  $\text{C}_{10}$  and  $\text{C}_{15}$  that match the more prominent series observed in the chromatogram (Fig. 3.4). The reduced abundance of  $\text{C}_{12}\text{H}_{22}$  products (less red) highlights the improbability of breaking multiple carbon-carbon bonds during degradation. The large backbone size of carotenoids ( $\text{C}_{40}$ ) would easily accommodate the wide size range of observed hydrocarbon products from  $\text{C}_{10}$  to  $>\text{C}_{25}$  (Fig. 3.4). The reduction process does not add or remove carbon atoms, and so, this size diversity is present in native DOM. The isomeric diversity of products, observed as many peaks of identical mass within a region, also unambiguously confirms previously implied isomerism in native DOM samples (Hertkorn et al 2013, Dittmar 2015). Consistent with the highly unsaturated tail of carotenoid molecules is the remarkable isomeric diversity of reduction products, a trait

previously attributed to PPL-DOM (Witt et al. 2009) but never directly detected. Using molecular ions ( $M^+$ ) detected for several compounds it was possible to define regions of the chromatogram that contained compositionally related molecules. For example, all the peaks within the  $C_{12}H_{22}$  region, indicated by a dotted ellipse, contain similar mass spectra with identical molecular ions (Fig. 3.4). The regions to the left and right are offset (-/+ ) by 14 Da (a difference of  $CH_2$ ), resulting in isomers of the chemical formula  $C_{11}H_{20}$  and  $C_{13}H_{24}$  respectively. The multiply unsaturated and branched carotenoid cleavage products (Fig. 3.4) must be highly oxidized at different locations to be consistent with NMR data, which could ultimately lead to some rearrangement of the carotenoid carbon skeleton either in native DOM or during the reduction. Previously, the reduction mechanism was shown to produce isomers from certain types of model compounds (Chart 1 and Fig. 1 in Arakawa & Aluwihare 2015). However, the large chromatographic regions over which isomers elute confirm the presence of multiple isomers in native DOM samples (i.e., progressively larger contiguous carbon skeletons that cannot be an artifact of the reduction).

The recovered 10% of the carbon post-reduction, while significant, is a small fraction of the original sample. The reduction process has poor recoveries of certain model compounds, in particular those with olefinic protons (Arakawa & Aluwihare 2015). Thus, the observed yield certainly underestimates the true percentage of degraded carotenoid material in DOM. It was therefore necessary to seek another method, in this

case, NMR, to accurately interrogate the anticipated abundance of degraded carotenoids in DOM.

Terpenoids, and specifically carotenoids, as known precursors to DOM were first proposed using advanced NMR techniques (Leenheer et al. 2003, Hertkorn et al. 2006, Lam et al. 2007, Woods et al. 2012). Past studies applied heteronuclear multiple bond correlation experiments (HMBC) to identify the major fragments in DOM but combining fragments to provide complete structures was not possible (Lam et al. 2007). However, GC-MS identification of carotenoid degradation products provides a definitive model to augment previous interpretations of NMR data. Here, Heteronuclear Single Quantum Coherence experiments (HSQC; Fig. 3.5A and 3.5C), which correlate proton chemical shifts and carbon chemical shifts over one bond in a sample, were combined with HMBC (Fig. 3.5B and 3.5D), which correlates long-range functional group interactions. Together HSQC units identify the H-C fragments in a structure and HMBC explains how these H-C fragments link together to form a complete molecule.

The GC-MS identification of a size range of degraded carotenoids with alicyclic head groups complements NMR data which bears the fingerprint of relatively intact carotenoid structures. In the HMBC spectrum it is possible to pick out both key one bond correlations from the conjugated double bonds (orange circle, Fig. 3.5) as well as the long range correlation between the double bond and the methyl units (blue dashed circle, Fig. 3.5D), which are distributed throughout carotenoid structures. The abundance of methyl functionalities in PPL-DOM is demonstrated by the edited HSQC (Fig. 3.5C, green and purple highlights).

However intact carotenoids have very low aqueous solubility, suggesting the majority of molecules must have been degraded/transformed to make them polar enough to become water-soluble (Fig. 3.6A vs. 3.6C). Generating hypothetical oxidized carotenoid structures consistent with the NMR data was challenging; spectra of 100's of potential structures containing hydroxyl (OH) and carboxyl (COOH) functionalities were simulated to identify matches to PPL-DOM resonances. These simulations indicated that OH groups are unlikely to be present on the terminal ring (no prominent long range CH<sub>2</sub>-CHOH couplings were detected; Fig. 3.5D vs. Fig. 3.6B, region V). Therefore, the heavily unsaturated main chain unit seems to be the logical point to introduce much of the oxidation necessary to satisfy NMR correlations, and indeed these double bonds are likely the most reactive sub-unit (Britton 1995). Interestingly, only NMR simulations where model structures contain fully hydroxylated chains accurately approximate the dominant correlations observed in PPL DOM (Fig. 3.6B). Considering that COO<sup>-</sup> (44Da) losses are commonly observed when PPL-DOM is analyzed by mass spectrometry (Reemtsma et al. 2008), it is likely that carboxylation of the structure has also taken place. Position-specific enzymatic cleavage of the main chain can produce ketones and aldehydes that are subsequently oxidized to carboxylic acids (Schwartz et al. 1997). Additionally, methyl groups may also be converted into COOH groups (Kitaoka et al. 2015). For the purpose of reproducing observed NMR correlations the random carboxylation of a portion of the methyl groups produces correlations that are then consistent with DOM (Fig. 3.7B and 3.6C).

Modeled NMR correlations suggest that carotenoids and oxidized carotenoids are abundant in PPL-DOM. Absolute quantification of carotenoid-like components is

challenging due to spectral overlap with other compounds. However, these components clearly dominate the HMBC experiment. Considering HMBC is one of the most insensitive NMR experiments the detection of these structural components during this experiment suggests they are amongst the most abundant in PPL-DOM. Estimations for both HSQC and carbon data suggest these regions may account for approximately 50% of the total signal. Signals from relatively intact linear carotenoids (i.e., conjugated double bonds) are ~10 times less intense when compared to oxidized components in the HMBC, suggesting that the oxidized components dominate in the mixture.

The PPL-DOM has a measured radiocarbon signature of -183‰ (Table 3.1) corresponding to a  $^{14}\text{C}$  age of 1625 years, which is consistent with previous reports for marine PPL-DOM (Lechtenfeld et al. 2014, Flerus et al. 2012). Upon acid hydrolysis of PPL-DOM it was noted that several small polar compounds (sugars, amino acids, and small acids) were liberated that could confound the potential, true radiocarbon signature of the terpenoid fraction. Isolation of the terpenoid fraction from these compounds was achieved by a second SPE step, which yielded carotenoid-rich DOM with a measured radiocarbon value of -240‰, or 2700 radiocarbon years (Table 3.1, Fraction 1).  $^1\text{H}$ NMR experiments showed that this fraction is rich in the methyl proton resonances characteristic of intact and oxidized carotenoids (Fig. 3.8); a carbon mass balance revealed it to represent 58% of PPL-DOM carbon, or approximately 23% of the total ambient dissolved organic carbon (DOC). The radiocarbon value of -240‰ is still relatively enriched compared to deep ocean DOM (e.g., -540‰) (Williams & Druffel 1987). Therefore, carotenoid degradation products while contributing to refractory DOM in the surface ocean, must also occupy a range of  $^{14}\text{C}$  signatures in surface waters. This is

the first study to link the structural components previously attributed to refractory DOM (Hertkorn et al. 2006, Lam et al. 2007) to a depleted radiocarbon signature.

The findings reported here support post biosynthesis modification of abundant biochemicals -carotenoids- to produce the isomeric diversity observed for PPL-DOM (Fig. 3.9). That carotenoid-related material is structurally distinct from the biosynthetic precursor could support the hypothesis that it is intrinsically refractory. Perhaps most important is the assignment of structure and potential biosynthetic precursors to a component of refractory DOM in aquatic systems. This is a crucial step forward as it enables testable hypotheses to be designed to identify important processes (e.g., photochemistry versus the microbial carbon pump) that shape refractory DOM, and control its removal from the ocean (Lang et al. 2006, Hawkes et al. 2015). Carotenoids are widespread in marine photosynthetic organisms, evidenced by the variety observed in marine particulate organic matter (Repeta & Gagosian 1982). Therefore, data presented here links the ocean's Biological Carbon Pump to DOM and provides evidence for the long-term accumulation of a product of marine photosynthesis in the DOM reservoir.

### **Acknowledgements**

Chapter III, in part, has been submitted for publication of the material as it may appear in *Science Advances*, 2016, Arakawa, Neal; Lane-Coplen, Daniel; Soong, Ronald; Stephens, Brandon; Simpson, Andre J.; Aluwihare, Lihini. The dissertation author was the primary investigator and author of this paper.

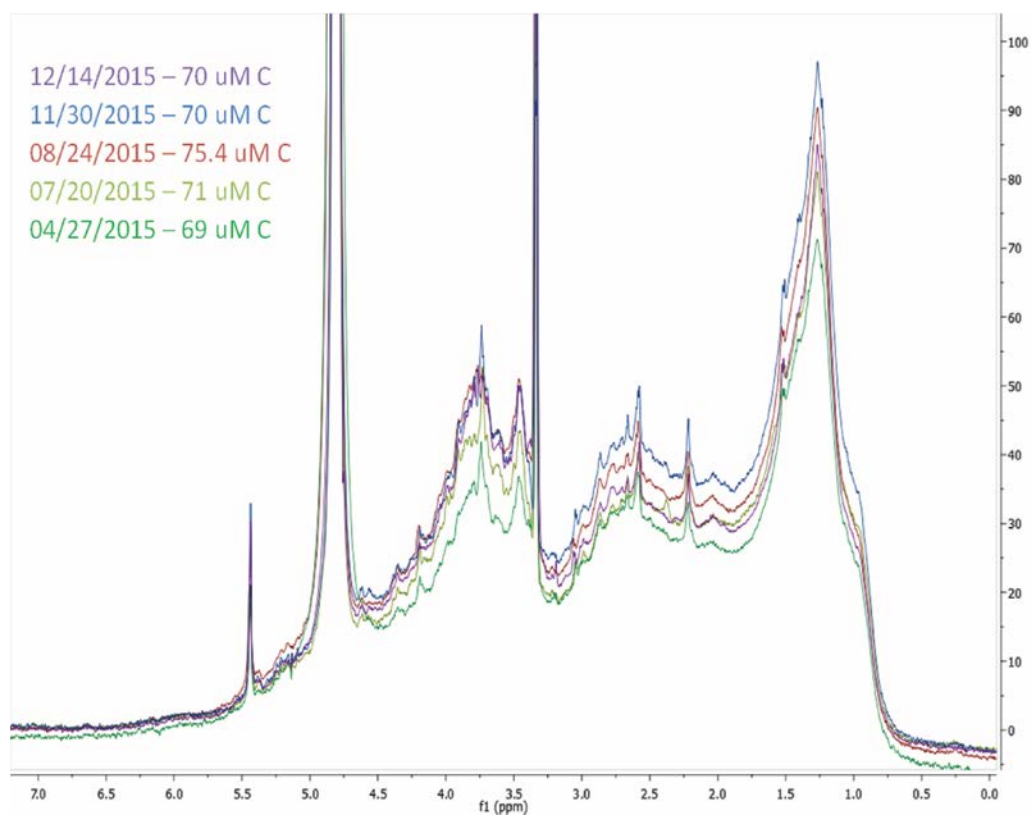


Figure. 3.1.  $^1\text{H}$  NMR spectra of 7 PPL-DOM, time series samples collected at the SIO Pier



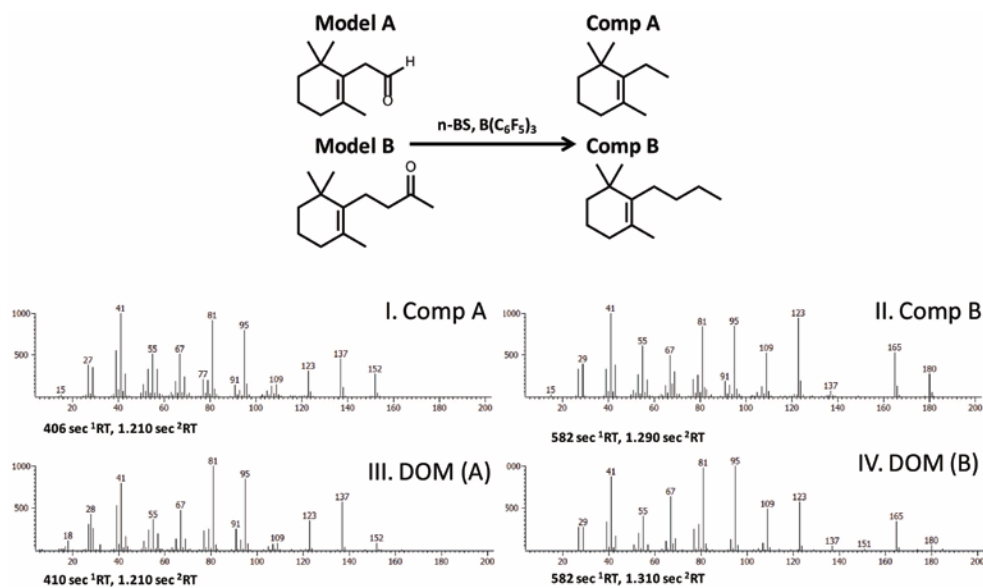


Figure. 3.2. Structures given for model compounds (Model A, Model B), and resulting reduced hydrocarbons (Comp A, Comp B). I, II. Mass spectra for Comp A and Comp B are consistent prominent peaks identified as III, IV (DOM A, B from Fig. 3.4). Retention times in Dimension 1 (<sup>1</sup>RT) and Dimension 2 (<sup>2</sup>RT) are also consistent.

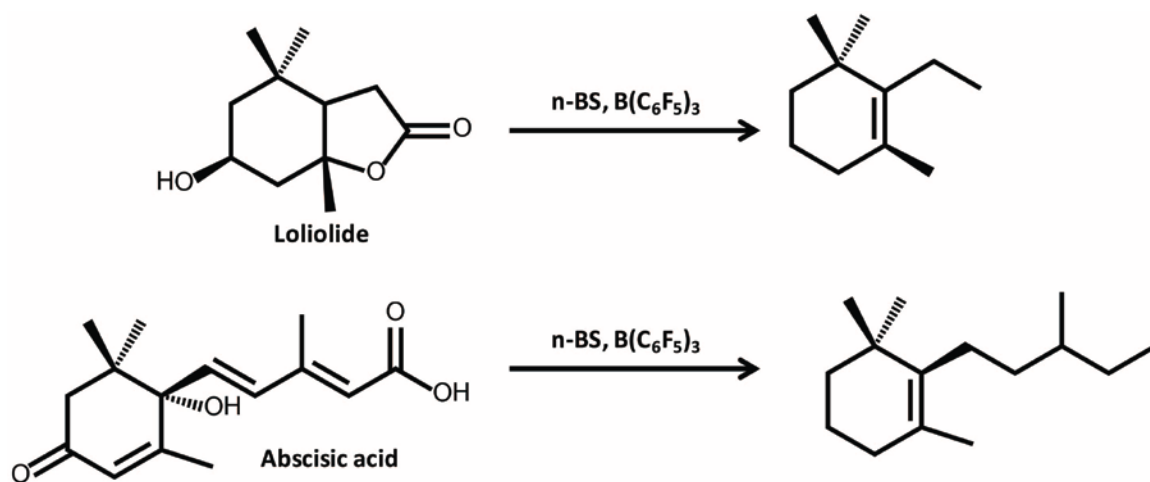


Figure. 3.3. Structures for known carotenoid degradation products Loliolide and Abscisic acid, and potential resulting reduced hydrocarbons.

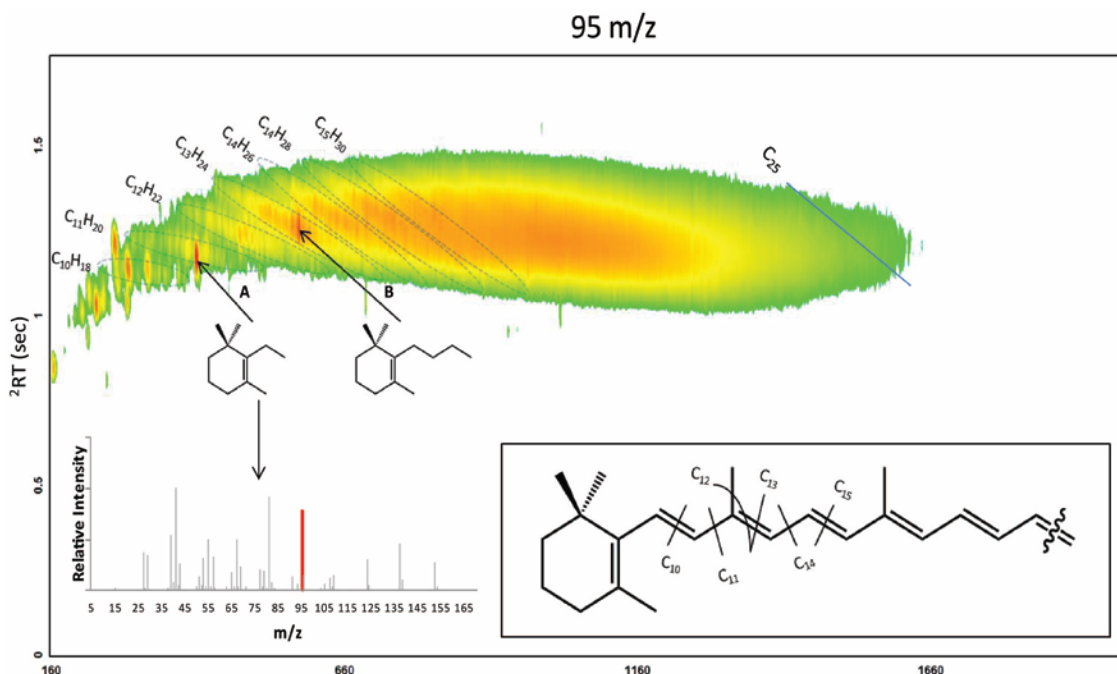


Figure. 3.4. GCxGC selected-ion chromatogram ( $95\ m/z$ ) of chemically reduced, PPL-extracted marine DOM (PPL-DOM). Ion abundance displayed as green (low)  $\Rightarrow$  red (high). Individual compounds appear as sharp ellipses. Exact chemical structures are linked to two distinct peaks (A, B), and the mass spectrum of A is provided with  $m/z\ 95$  - a prominent hydrocarbon fragmentation ion - identified in red. Chemical formulas shown in the figure were validated with hydrocarbon standards. Inset shows a generic carbon backbone of carotenes/carotenoids and identifies positions where carbon bond cleavage can form alicyclic compounds of sizes ( $C_{10}$ - $C_{15}$ ) denoted in the chromatogram. Blue line estimates the range of  $C_{25}$  hydrocarbons, which are among the largest compounds identified in marine PPL-DOM.

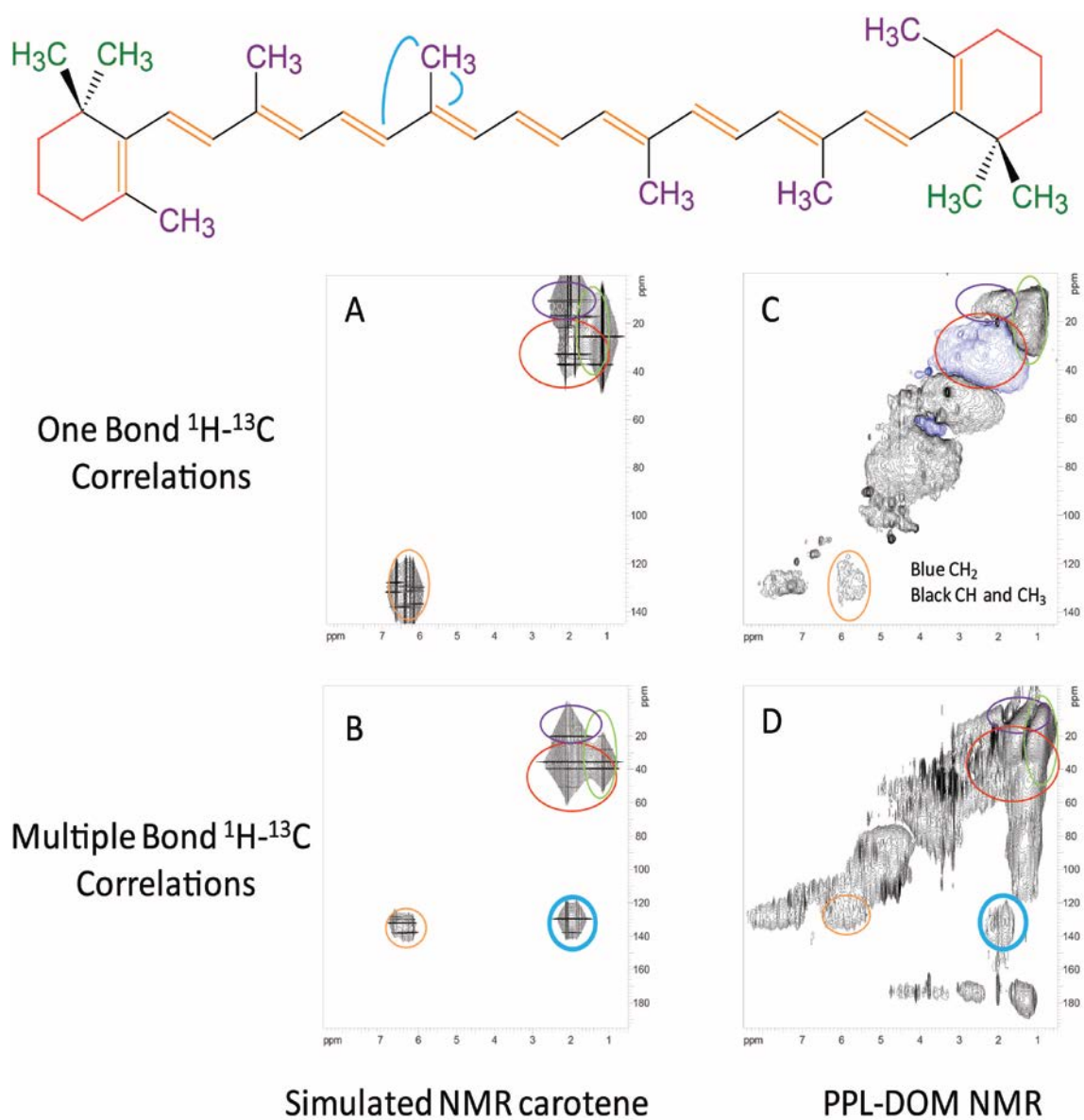


Figure. 3.5. Identification of intact carotenoids through  $^1\text{H}$ - $^{13}\text{C}$  NMR correlations. One bond (A), and multiple bond (B) modelled  $^1\text{H}$ - $^{13}\text{C}$  correlations for an intact carotene. Colors highlighted on the structure correspond to the same colored regions in each spectrum. One bond (C) and multiple bond (D)  $^1\text{H}$ - $^{13}\text{C}$  correlations for PPL-DOM, with overlaid regions. Blue regions correspond to long-range correlations.

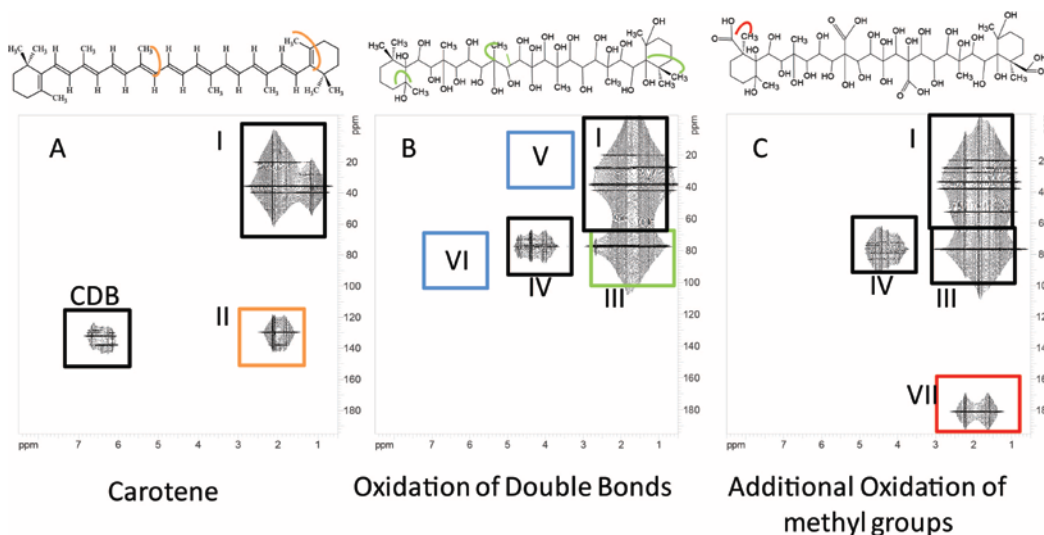


Figure. 3.6. (A). Long range  $^1\text{H}$ - $^{13}\text{C}$  correlations for carotene. Region I shows terminal ring alkane correlations, CDB and II show correlations to conjugated double bonds (orange in structure). (B). Long range  $^1\text{H}$ - $^{13}\text{C}$  correlations for hydroxylated carotene. Regions III and IV identify correlations to COH groups. Region V is where correlations between COH and ring alkyl groups would be expected if the terminal ring is hydroxylated. Region VI is where coupling between COH and double bonds would be expected if the main chain was only partially hydroxylated. (C). Long range  $^1\text{H}$ - $^{13}\text{C}$  correlations between hydroxylated and partially carboxylated carotene (red, region VII). Black, green, orange and red boxes correspond to similarly color-coded NMR correlations and are simulated. Blue boxes identifying missing resonances as described in the text. The intact, fully oxidized structures presented are used to generate model NMR data. However, as confirmed by GC data, a complex mixture of smaller oxygenated products with different chain lengths is likely more representative of the DOM sample.



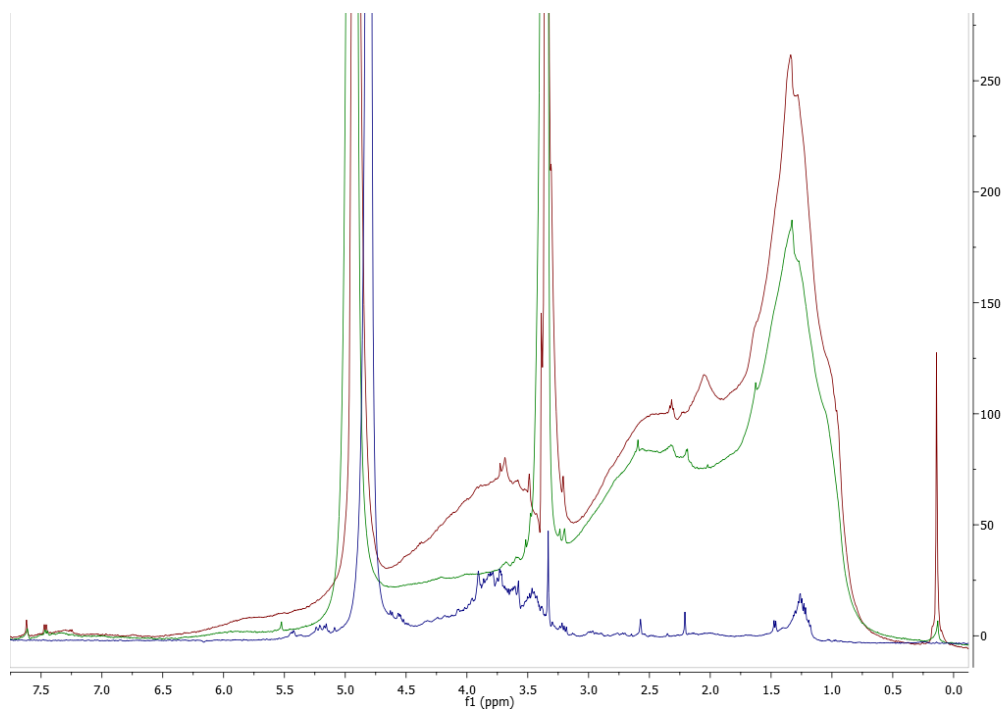


Figure. 3.8.  $^1\text{H}$  NMR spectra of PPL-DOM (Red), Carotenoid-rich fraction (Green; retained on 2<sup>nd</sup> PPL column following acid hydrolysis, see text), and sugar and amino acid-rich fraction (Blue; not retained by 2<sup>nd</sup> PPL column following acid hydrolysis).

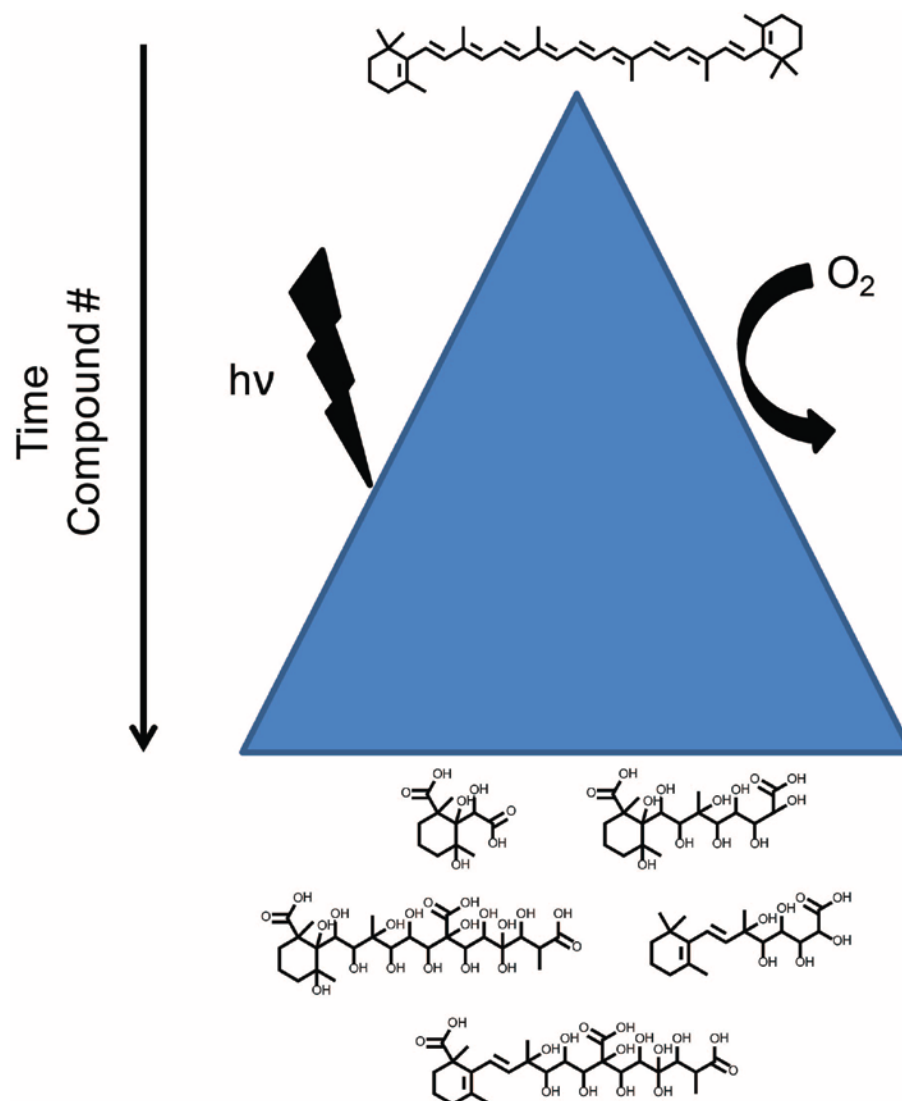


Figure. 3.9. Conceptual figure depicting how carotenoid degradation contributes soluble, structurally diverse molecules to the DOM reservoir. Photochemistry ( $h\nu$ ) and other oxidative processes ( $O_2$ ) are depicted as leading mechanisms of post-biosynthesis modification,



Table 3.1. Isotope and elemental data for PPL-DOM samples collected from SIO Pier (32.87°N, 117.26°W). At this location typical total DOC concentrations range from 70-80  $\mu$  M and chlorophyll *a* is between 1-2  $\mu$ g L<sup>-1</sup>. PPL-DOM and carotenoid-rich fractions have  $\delta^{13}\text{C}$  values that reflect those of 16 temporally spaced PPL-DOM samples obtained from a monthly SIO Pier time series ( $-23.1 \pm 0.6\%$ ).

	$\Delta^{14}\text{C}$ (‰)	$\delta^{13}\text{C}$ (‰)	$\delta^{15}\text{N}$ (‰)	C/N
PPL-DOM	-183	-21.8	4.3	26
Carotenoid-Rich	-286	-22.1	4	27.7

## References

- Arakawa, N., Aluwihare. 2015. Direct Identification of diverse alicyclic terpenoids in Suwannee River Fulvic Acid. *Environmental Science and Technology* 49: 4097-4105.
- Britton, G. 1995. Structure and properties of carotenoids in relation to function. *FASEB J.* 9: 1551-1558.
- Cicero, D.O, Barbato, G., Bazzo, R. 2001. Sensitivity enhancement of a two-dimensional experiment for the measurement of heteronuclear long-range coupling constants, by a new scheme of coherence selection by gradients. *Journal of Magnetic Resonance* 1: 209-213.
- Dittmar, T., Koch, B., Hertkorn, N., Kattner, G. 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnology and Oceanography: Methods*: 6: 230-235
- Dittmar, T. 2015. Reasons behind the long-term stability of dissolved organic matter in *Biogeochemistry of Marine Dissolved Organic Matter 2<sup>nd</sup> edition*. Hansell, D.A., Carlson, C.A., eds. Elsevier: 369-388
- Druffel, E.R.M., Williams, P.M. 1987. Radiocarbon in dissolved organic matter in the central North Pacific Ocean. *Nature* 330: 246-248.
- Flerus, R., Lechtenfeld, O.J. Koch., B.P., McCallister, S.L., Schmitt-Kopplin, P., Benner, R., Kaier, K., Kattner, G. 2012. A molecular perspective on the ageing of marine dissolved organic matter. *Biogeosciences* 9: 1935-1955
- Hawkes, J.A., Rossel, P.E., Stubbins, A., Butterfield, D., Connelly, D.P, Achterberg, E.P., Koschinsky, A., Chavagnac, V., Hansen, C.T., Bach, W., Dittmar, T. 2015. Efficient removal of recalcitrant deep-ocean dissolved organic matter during hydrothermal circulation. *Nature* 7: 51-54.
- Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., Kettrup, A., Hedges, J.I. 2006. Characterization of a major refractory component of marine dissolved organic matter. *Geochim Cosmochim Acta* 70: 2990-3010.
- Hertkorn, N., Harir, M., Koch, B.P., Michalke, B., Schmitt-Kopplin, P. 2013. High-field NMR spectroscopy and FTICRMS mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosciences* 10: 1583-1624.
- Kitaoka, N., Lu, X., Yang, B., Peters, R.J. 2015. The application of synthetic biology to elucidation of plant mono-, sesqui-, and diterpenoid metabolism. *Molecular Plant* 8: 6-16.

- Lam, M., Baer, A., Alae, M., Lefebvre, B., Moser, A., Williams, A., Simpson, A.J. 2007. Major structural components in freshwater dissolved organic matter. *Environmental Science and Technology* 41: 9240-9247
- Lang, S.Q., Butterfield, D.A., Lilley, M.D., Jognson, H.P., Hedges, J.I. 2006. Dissolved organic carbon in ridge-flank and ridge-axis environments. *Geochim. Cosmochim. Acta* 70: 3830-3842.
- Lechtenfeld, O. J., Kattner, G., Flerus, R., McCallister, S. L., Schmirrt-Kopplin, P., Koch, B. P. 2014. Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean. *Geochim. Cosmochim. Acta* 126: 321-327
- Leenheer, J.A., Nanny, M.A., McIntyre, C. 2003. Terpenoids as major precursors of dissolved organic matter in landfill leachates, surface water, and groundwater. *Environmental Science and Technology* 37: 2323-2331.
- Liaaen-Jensen. 1991. Marine carotenoids: recent progress. *Pure & Applied Chemistry* 63: 1-12.
- Nimmagadda, R.D., McRae, C.R. 2007. Characterisation of the backbone structures of several fulvic acids using a novel selective chemical reduction method. *Organic Geochemistry* 38: 1061-1072.
- Reemtsma, T., These, A., Linscheid, M., Leenheer, J., Spitzzy, A. 2008. Molecular and structural characterization of dissolved organic matter from the deep ocean by FTICR-MS, including hydrophilic nitrogenous organic molecules. *Environmental Science and Technology* 42: 1430-1437.
- Repeta, D.J., Gagosian, R.B. 1982. Carotenoid transformation in coastal marine waters. *Nature* 295: 551-54.
- Repeta, D.J., Aluwihare, L.I. 2006. Radiocarbon analysis of neutral sugars in high-molecular-weight dissolved organic carbon: Implications for organic carbon cycling. *Limnol. Oceanogr.* 51: 1045-1053.
- Schwartz, S.H., Tan, B.C., Gage, D.A., Zeevaart, J.A., McCarty, D.R. 1997. Specific oxidative cleavage of carotenoids by VP14 of maize. *Science* 276: 1872-1974.
- Sexton, P.F., Norris, R.D., Wilson, P.A., Palike, H., Westerhold, T., Rohl, U., Bolton, C.T., Gibbs, S. 2011. Eocene global warming events driven by ventilation of oceanic dissolved organic carbon. *Nature* 471: 349-352.
- Siegenthaler, U., Sarmiento, J.L. 1993. Atmospheric carbon dioxide and the ocean. *Nature* 365: 119-125.

- Simpson, A.J., McNally, D.J., Simpson, M.J. 2011. NMR spectroscopy in environmental research: from molecular interactions to global processes. *Prog Nucl Magn Reson Spectrosc.* 58: 97-175.
- Stenson, A.C., Marshall, A.G., Cooper, W.T. 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: 1275-1284.
- Stuiver, M., Polach, H.A. 1977. Discussion: reporting of  $^{14}\text{C}$  data. *Radiocarbon* 19: 355-363.
- Taylor, H.F., Burden, R.S. 1970. Identification of plant growth inhibitors produced by photolysis of violaxanthin. *Phytochemistry* 9: 2217-2223.
- Witt, M., Fuchser, J., Koch, B.P. 2009. Fragmentation studies of fulvic acids using collision induced dissociation fourier transform ion cyclotron resonance mass spectrometry. *Analytical Chemistry* 81: 2688-2694.
- Woods, G.C., Simpson, M.J., Simpson, A. 2012. Oxidized sterols as a significant component of dissolved organic matter: Evidence from 2D HPLC in combination with 2D and 3D NMR spectroscopy. *Water Research*: 46: 3398-3408.

## **Chapter IV**

Oxidation of  $\beta$ -carotene produces compounds that resemble dissolved organic matter

## Abstract

We oxidized  $\beta$ -carotene in both freshwater (MilliQ) and filtered seawater, recovering up to 6.6% of the original carbon as a mixture of complex water-soluble oxidation products. Degraded compounds have a proton ( $^1\text{H}$ ) nuclear magnetic resonance (NMR) spectrum resembling that of marine dissolved organic matter (SIO DOM). Analysis by comprehensive gas chromatography (GCxGC) time of flight (TOF) mass spectrometry (MS) confirmed a large compositional diversity of degraded compounds, some of which share mass spectra similar to compounds found in SIO DOM. Finally, chemical reduction of degradation products resulted in a series of alicyclic hydrocarbons also observed in reduced SIO DOM, and previously attributed to degraded carotenoids, confirming these ubiquitous compounds as a source of a significant fraction of DOM that accumulates in aquatic environments.

## Introduction

Carotenoids are organic biomolecules which function as light harvesting pigments in photosynthesizing organisms, and are therefore ubiquitous in nature. Generally, they exist as  $\text{C}_{40}$  compounds with an extended network of conjugated double bonds, synthesized from 5-carbon branched (isoprene) building blocks (Liaen-Jensen, 1991). The contribution of carotenoids to either terrestrial or marine dissolved organic matter (DOM) is understudied; although carotenoids have been widely observed in sediments and particulate organic matter (Repeta & Gagosian, 1982, 1987), only lately have structural features indicative of carotenoids been observed in DOM (Lam et. al., 2007). Recently, degraded carotenoids were directly detected in a marine DOM sample (Chapter

3), yet there is no model for DOM formation from carotenoids. Here, we demonstrate that structurally diverse compounds that closely resemble major DOM structural characteristics can be produced through abiotic oxidation of  $\beta$ -carotene, a single carotenoid. There is much debate over what drives compositional diversity in DOM (Dittmar, 2015), and our results demonstrate the potential significance of abiotic processes. Furthermore, this represents the first study to observe considerable molecular level transformation of water-insoluble biomolecules into dissolved compounds, all over relatively short timescales. Such processes are rarely considered to significantly contribute to DOM formation (Hansell and Carlson 2015). We expect this work to be a starting point for a wide range of studies, from mechanistic insights into carotenoid oxidation to high-resolution mass spectrometric targeting of carotenoid degradation products within environmental DOM samples.

## **Materials and Methods**

### *Oxidation of $\beta$ -carotene and isolation of products*

$\beta$ -Carotene (BC) was purchased from Sigma Aldrich and stored at  $-20^{\circ}\text{C}$ . For oxidations, 20-60 mg of BC was placed into combusted quartz tubes. Either MilliQ water, or pre-filtered (AcroPak™ 0.8/0.2  $\mu\text{m}$  polyethersulfone (PES) Supor® membrane filter) seawater from the Scripps Institution of Oceanography (SIO) Pier (UCSD, La Jolla) was added to the tubes. “Dark” replicates were covered with aluminum foil and/or opaque plastic bags. All tubes were placed in a seawater bath on the SIO Pier, and left exposed to natural sunlight cycles. To extract samples, tube contents were emptied into separation funnel and tube was rinsed with dichloromethane (DCM). DCM layer was

extracted from aqueous layer (1:1 ratio). Aqueous layer was extracted again 2x with DCM, and organic layers were combined as Residual starting material. Emulsion layer was collected (if present) , and the aqueous layer was bubbled with N<sub>2</sub> gas under mild heat (50°C) for 30 minutes to remove residual DCM. Aqueous layer was then acidified to pH ~2 with concentrated HCl. Aqueous-soluble BC degradation products were then isolated by extraction onto 1 gm Agilent Bond Elut PPL styrene-divinylbenzne polymer cartridges @ ~15ml/min. Prior to elution, cartridges were rinsed with 2 cartridge volumes 0.01 M HCl, followed by 2 cartridge volumes pure water (MilliQ). The cartridges were then dried under N<sub>2</sub> gas, and eluted with 2 cartridge volumes of MeOH. The combined eluents were dried under N<sub>2</sub> gas, resulting in PPL-extractable.

#### *DOM isolation from SIO Pier*

The isolation protocol was adapted from Dittmar et al. (2008). 330L of seawater from the Scripps Institution of Oceanography Pier (UCSD, La Jolla), were collected and pre-filtered through an AcroPak™ 0.8/0.2 µm polyethersufone (PES) Supor® membrane filter. The filtrate was acidified to pH 2 with conc. Hydrochloric acid (HCl), and passed over activated (1 cartridge volume methanol (MeOH)) 1 gm Agilent Bond Elut PPL styrene-divinylbenzne polymer cartridges @ ~15ml/min. 9 cartridges were used and eluted after 20 L. Prior to elution, cartridges were rinsed with 2 cartridge volumes 0.01 M HCl, followed by 2 cartridge volumes pure water (MilliQ). The MilliQ rinse was done to limit methylation of DOM under MeOH elution. The cartridges were then dried under N<sub>2</sub> gas, and eluted with 2 cartridge volumes of MeOH. The combined eluents were dried



under N<sub>2</sub> gas, resuspended in MilliQ, frozen, and lyophilized. The total weight of recovered DOM was 235 mg.

#### *Elemental and Stable Isotope (<sup>13</sup>C ) Analysis*

Elemental measurements were made at the Scripps Institution of Oceanography using standard elemental analyzer isotope ratio mass spectrometry (EA-IRMS).

Elemental carbon is reported as %C. Nitrogen content of all samples was extremely low (0.01-0.2%). Stable isotopes of carbon is reported in standard δ notation (‰, per mil) relative to Pee Dee Belemnite (PDB).

#### *Proton (<sup>1</sup>H) Nuclear Magnetic Resonance (NMR) Spectroscopy*

Samples (1-4 mg) were dissolved in 0.5 mL of deuterated methanol (CD<sub>3</sub>OD) or deuterated chloroform (CDCl<sub>3</sub>) and <sup>1</sup>H-NMR spectra were determined on a 500MHz Varian NMR spectrometer at the Scripps Institution of Oceanography. Typically, samples were typically acquired with 128 scans. NMR spectra were referenced to either 7.26 ppm (CDCl<sub>3</sub>) or 3.31 ppm (MeOD). Mestrenova software was used to integrate spectra.

#### *Derivatization procedure*

Samples were weighed (~1mg, in MeOH, dried overnight at 60°C) into combusted 400 μL GC Vial inserts on a balance accurate to ±0.05 mg. Samples were derivatized in 400μL (3:1 BSTFA10%TMCS: Pyridine, Ball and Aluwihare 2014). GC vials were capped and heated at 70°C for 1 hour. Samples were then analyzed by GC-MS/FID and GCxGC-TOF-MS

### *GC-MS/FID Analysis*

An Agilent 7890A Gas Chromatograph system coupled simultaneously to an Agilent 5975C quadrupole mass spectrometer and a flame ionization detector was used for GC-MS/FID characterization of the sample. Splitless injection with 1  $\mu\text{L}$  of analyte was used. Separation was performed on a 5% phenyl poly(dimethylsiloxane) column (J&W 123-5731DB-ht, 30 m, 320  $\mu\text{m}$  i.d., 0.1  $\mu\text{m}$  film) with a temperature program from 100  $^{\circ}\text{C}$  (hold time 2 min) to 200  $^{\circ}\text{C}$  (at 5  $^{\circ}\text{C}/\text{min}$ , hold time 0 min) to 320  $^{\circ}\text{C}$  (at 8  $^{\circ}\text{C}/\text{min}$ , hold time 8 min). Helium was the carrier gas with a constant flow of 1.8 mL/min. After separation, the effluent was split between the flame ionization detector (operating at 310 $^{\circ}\text{C}$ ) and the mass spectrometer (70eV ionization, scanning 50-750 m/z). Each FID signal was manually integrated in triplicate and the average value was used. The average standard deviation for triplicates was 0.6%. FID areas are either presented normalized to sample amount, or by percent FID values corresponding to each region of interest (where percent was calculated by dividing FID area of interest by total FID area).

### *Comprehensive Two-Dimensional Gas Chromatography (GCxGC-TOF-MS)*

The GCxGC instrument used is a LECO Pegasus 4D GCxGC-Time-of-Flight Mass Spectrometer (GCxGC-TOF-MS). The term GCxGC refers to the use of two distinct columns in series that have different chemical selectivity. Compounds are separated primarily by volatility in the first column and polarity in the second. The analysis is comprehensive because all of the effluent from the first column is cryofocused and transferred onto the second column. Effluent from the second column is then analyzed by the TOF-MS, which benefits from high spectral acquisition speed (50 Hz-500 Hz).

Finally, the data is compiled into a two-dimensional chromatogram that is visualized and processed by ChromaTOF® software. Both columns are housed within an Agilent 7890A Gas Chromatograph. The splitless inlet temperature is set at 300°C. The first dimension column is a semi-polar Crossbond® diphenyl dimethyl polysiloxane column (Restek Rxi-17Si, 30m length, 0.2550mm i.d., 0.25µm film thickness). For analysis of reduced products, the column was programmed to remain isothermal at 40°C for 1 min, and ramped to 315°C at 3°C per minute. The modulator temperature was offset by +30°C to the primary oven. The secondary oven (within the GC) housed the second dimension non polar Crossbond® dimethyl polysiloxane column (Restek Rxi-1, 1.58m length, 0.250mm i.d., 0.25µm film thickness). The secondary oven temperature was offset by +25°C to the primary oven. The modulation period was 2.5 s, with a hot pulse time of 1.05 s and a cool time of 0.2 s. For analysis of TMS-derivatized compounds, temperature program was as follows; 115°C for 3 min, 320°C at 8°C/min. Modulation period of 4s, 1.6s hot pulse time, 0.4s cool time. The carrier gas was Helium at a constant flow of 1.5 mL/min. The acquisition delay on the TOF-MS was set to 160s. The acquisition rate was set to 50Hz, with a range of 5-1000 m/z. Electron Ionization was run at 70eV.

#### *Reduction procedure*

The reduction procedure directly follows Arakawa and Aluwihare (2015) which was modeled after Nimagadda et al. (2007). For reductions, compounds (~5 mg) were weighed to a flame-dried 2 mL vial equipped with a stir bar. Under Argon atmosphere, 10 mg of  $B(C_6F_5)_3$  was added (100 µL of 100 mg  $B(C_6F_5)_3$ /1 mL dichloromethane

solution). Immediately after addition, 100  $\mu\text{L}$  of n-BS was added, also under Argon. The mixture was then allowed to stir-overnight. Note that the sample is completely soluble in dichloromethane after the reduction. Following reduction the samples were treated to remove excess catalyst and siloxane by-products. Approximately 50  $\mu\text{L}$  of the reaction mixture was transferred into 250  $\mu\text{L}$  of conc. KOH/MeOH. After 1 minute the mixture was extracted with pentane (3x 100 $\mu\text{L}$ ). The pentane fractions were collected, washed with  $\text{H}_2\text{O}$ , and dried over  $\text{Na}_2\text{SO}_4$ . The organic fraction was then dried under  $\text{N}_2$  to 100 $\mu\text{L}$  for GC-FID/MS and GCxGC-TOF-MS analysis.

## Results and Discussion

Initially,  $\beta$ -Carotene was placed in a quartz tube with fresh water and exposed to natural light cycles. After 18 days the tube was processed to identify degradation products. Fractions recovered through processing have been labeled as residual starting material (Fig. 4.1B), emulsion (Fig. 4.1C), and PPL-extractable (Fig. 4.1D). Elemental analysis confirms that all fractions of recovered material are less carbon-rich than the starting material (91% C > 73% C; 61% C; 57% C, Fig. 4.1). A mass balance of these fractions found that 82% of the starting  $\beta$ -carotene carbon is recovered in the residual (unreacted) material fraction, while 2.9% is recovered in the emulsion fraction, and 2.7% recovered in the PPL-extractable fraction. Given the simple experimental conditions and low nitrogen content of all fractions (Table 4.1), the results demonstrate that all fractions of the starting material represent  $\beta$ -carotene compounds that are oxidized to varying degrees. The extent of oxidation has a direct effect on the solubility of degraded

products, such that the most oxidized compounds were observed in the PPL-extractable fraction. This fraction is directly comparable in isolation protocol to marine DOM samples and specifically to the SIO Pier PPL sample (Fig. 4.1E, and Chapter 3), isolated from surface waters near the Scripps Institution of Oceanography. Given that an initial extraction of  $\beta$ -carotene prior to any experimental manipulation did not show any water-soluble compounds, these results are interpreted to demonstrate direct oxidation of oxygen-free biomolecules to produce water soluble products.

NMR spectra of the isolated fractions largely support a model of progressive oxidation. Although the residual starting material fraction (Fig. 4.1B) appears largely similar to the starting material  $\beta$ -carotene (Fig. 4.1A), there is a slight increase in the ratio of alkyl (0.5-2.5 ppm) to allylic (5.5-7 ppm) protons from 3.27:1 to 4.57:1 ( $\beta$ -Carotene has a 3:1 ratio). This indicates a decrease in olefinic protons, likely due to oxidation and cleavage. Both the emulsion and PPL-extractable fractions contain comparatively small amounts of allylic protons. In conjunction with the elemental data and observed solubility, this is interpreted as further oxidation of double bonds. It is important to note, but not overemphasize, the striking visual similarities between the PPL-extractable fraction and SIO Pier samples. One notable feature present in both spectra is the broad alkyl region found between 1-1.3 ppm (Fig. 4.1D, 4.1E) associated with methyl groups in environmental DOM samples. Broad peaks such as this have been attributed to spectral overlap of innumerable individual compounds (references). This result is therefore surprising, given the simple, single compound starting material.

In order to examine the end production of the photooxidation, all fractions were derivatized and analyzed by GC-MS/FID and comprehensive GCxGC-TOF/MS. GC-MS/FID chromatograms support the oxidation/degradation of  $\beta$ -Carotene accompanied by a concurrent increase of small degradation products (Fig. 4.2). A few of these products have been identified as known carotenoid degradation products (Isoe et. al. 1969, Fig. 4.3). However, the emulsion and PPL-extractable fractions contain many unknown degradation products that are broadly distributed. Comprehensive GCxGC is a powerful tool for separating complex mixtures, accomplished through the use of two (chemically orthogonal) columns. The inclusion of an additional column creates an additional level of separation such that chromatograms are visualized as two-dimensional landscapes as opposed to a single chromatogram. Like GC-MS/FID chromatograms, the GCxGC chromatograms of all fractions (Fig. 4.4A-E) show the oxidation/degradation of  $\beta$ -carotene and the production of smaller degradation products.  $\beta$ -Carotene is present as both *cis* and *trans* isomers, thus leading to two different elution regions for the starting material (Fig. 4.4A, the large elution regions are largely due to column overloading). While  $\beta$ -carotene is present in residual starting material (Fig. 4.4B), it is not present in any other fraction. The chromatograms of the emulsion and PPL-extractable fractions (Fig. 4.4C, 4.4D) indicate that each fraction is a mixture of many different compounds in low individual concentrations. The presence of many structurally similar yet distinct compounds is consistent with the observed NMR peak broadening (Fig. 4.1C, D) for these samples. Comparison of the emulsion (Fig. 4.4C) and PPL-extractable fractions (Fig. 4.4D) also shows the preferential accumulation of more oxidized (and therefore larger – higher molecular weight and more polar compounds corresponding to later

elution times) degradation products in the PPL-extractable fraction. Finally, the chromatogram of the PPL-extractable fraction highlights the large diversity of oxidized degradation products, once again similar to the distribution of compounds observed for SIO Pier PPL-DOM (Fig. 4.4E). Material eluting in the same chromatographic regions of Fig. 4.4D and Fig. 4.4E possess similar mass spectra (Fig. 4.5). However, we cannot conclude from these data alone that the degraded compounds present in the PPL extract of oxidized  $\beta$ -carotene, are those that are observed in SIO Pier PPL-DOM. However, these initial results do confirm that once photooxidized,  $\beta$ -carotene can produce structurally diverse compounds that bear a close resemblance to compounds found in marine DOM (SIO Pier).

In a concurrent experiment, it was found that  $\beta$ -carotene oxidized in filtered seawater resulted in more PPL-extractable DOM (4%) than the Milli-Q experiment. Thus, in a second round of experiments, multiple bottles of  $\beta$ -Carotene in filtered seawater were exposed to natural sunlight cycles, and removed and processed at different times. After 7 days, the PPL-extractable fraction was recovered at 3.2% of initial carbon; at 24 days, 6.6% carbon; at 41 days, 6.3% carbon. For comparison a “dark” control recovered 2.7% of carbon at 42 days, and its PPL extractable material had a similar but not identical, NMR spectrum to its photochemically transformed counterparts (Fig. 4.6). These results are interpreted to suggest that oxidation processes are accelerated by light, but may occur whenever sufficient oxidation potential is available. The recovered material from filtered seawater oxidations was analyzed by NMR, and found to be largely similar to material recovered from fresh water trials (Fig. 4.7).

Photooxidized products from the  $\beta$ -carotene experiment conducted in filtered seawater were isolated by PPL and chemically reduced following previously established protocols (Chapter 2 and 3), and analyzed by GCxGC-TOF-MS. The most prominent feature observed in the reduction product was a series of unsaturated alicyclic hydrocarbons. Two of these compounds (peaks 1 and 3, Fig. 4.8A) were previously observed in the reduction products of marine DOM (Fig. 4.8B), and used as support for the presence of carotenoid degradation products. Peak 2 (Fig. 4.8A) is evenly spaced between 1 and 3, and contains a mass spectrum consistent with a C<sub>12</sub> alicyclic backbone of a degraded carotenoid (Fig. 4.9). This compound has never before been clearly and prominently observed in reduced environmental DOM samples. Furthermore, the mass spectra of other evenly spaced peaks (Fig. 4.9) displays new compounds in the series from C<sub>10</sub>-C<sub>14</sub>. Surprisingly, the distribution of reduction products is nearly identical to the distribution observed for reduced, marine DOM (Fig. 4.8B). The large size range of reduction products demonstrate that the  $\beta$ -carotene backbone is susceptible to oxidation at several points along its linear backbone, although smaller (<C<sub>20</sub>) reduction products are the most abundant. We have previously demonstrated that chromatographically resolvable isomeric diversity (diagonal curved regions in Fig. 4.8B; Chapter 2 and 3) of reduction products may be either present in native DOM samples or could be a function of the reduction process. However, the structural diversity (Peaks 1 vs. 2 vs. 3; Fig. 4.8A) is a real reflection of carbon backbone diversity in native DOM samples. Additionally, mass spectra of the chromatographically unresolvable areas of both reduced  $\beta$ -carotene degradation products and reduced marine DOM (Fig. 4.10) are highly similar. Thus, the



current study confirms previous observations implicating carotenoid degradation products as major components of terrestrial (Chapter 2) and marine (Chapter 3) DOM.

## **Conclusions**

Taken together, the results conclusively demonstrate that abiotic processes alone are sufficient to generate large amounts of structurally diverse carotenoid degradation products from a single carotenoid precursor. Thus given the 100s of known carotenoids from marine and terrestrial environments, it is easy to envision carotenoid degradation products as one driving force behind the incredible diversity of compounds found within environmental DOM samples.

## **Acknowledgements**

Chapter IV, in part is currently being prepared for submission for publication of the material. Arakawa, Neal; Aluwihare, Lihini. The dissertation author was the primary investigator and author of this material.

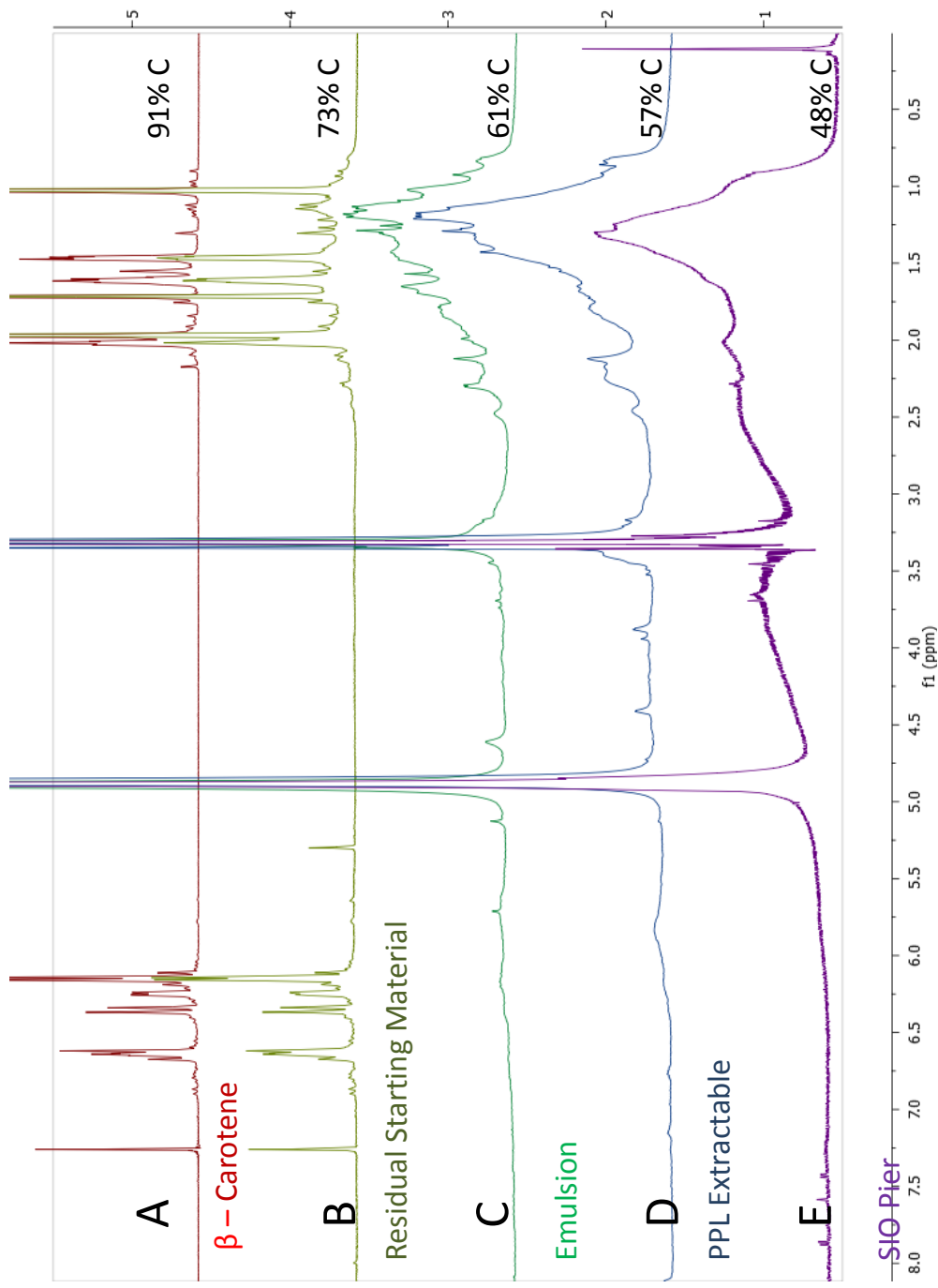


Figure 4.1. <sup>1</sup>H NMR spectra of  $\beta$ -carotene (A), and fractions of oxidized  $\beta$ -carotene products Residual starting material (B) Emulsion (C) and PPL Extractable (D). <sup>1</sup>H NMR spectrum of SIO Pier (E). Percent carbon for each fraction is presented on right side.

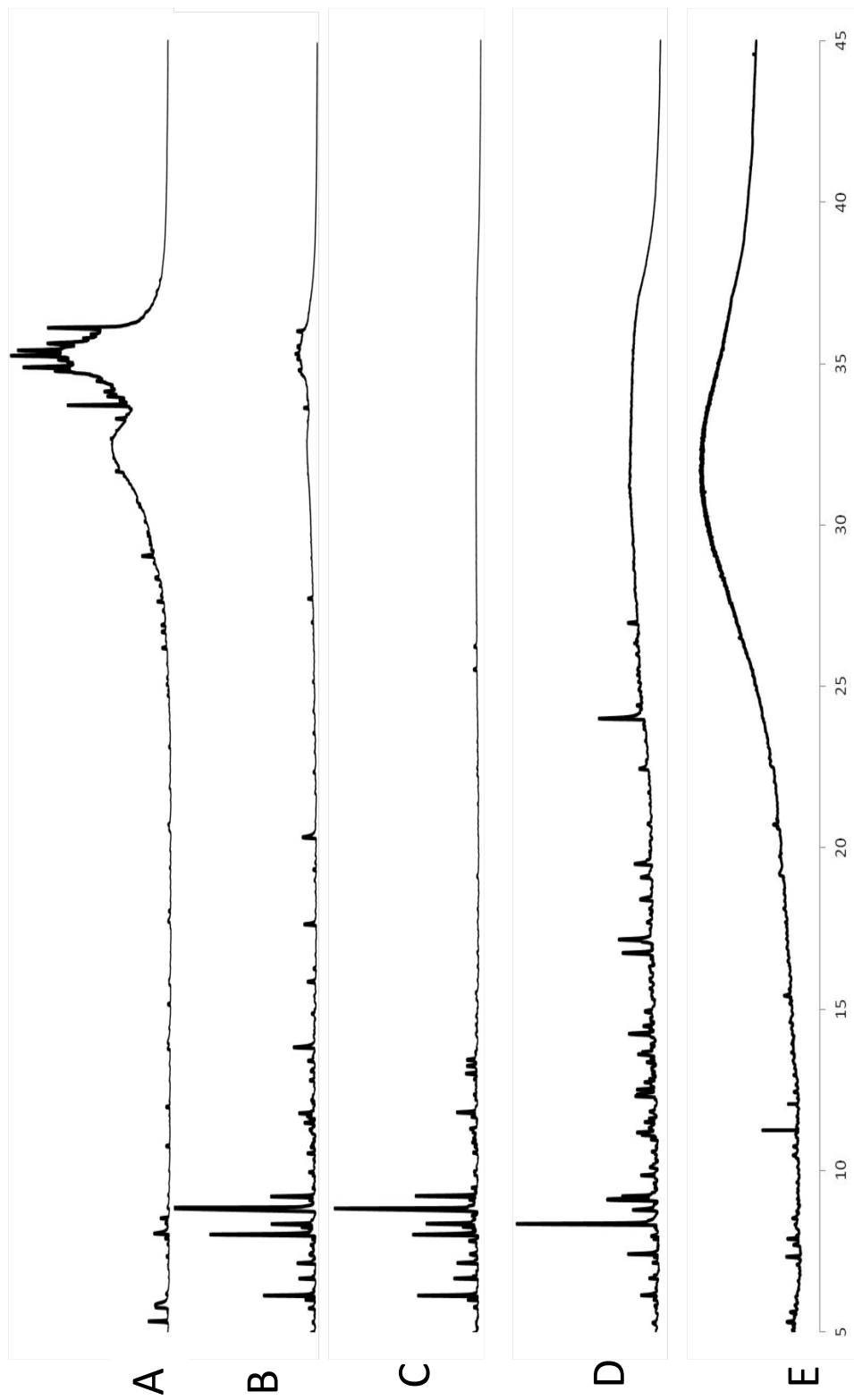


Figure 4.2. GC FID chromatograms of derivatized  $\beta$ -carotene (A), and oxidized  $\beta$ -carotene products Residual starting material (B) Emulsion (C) and PPL Extractable (D). Derivatized SIO Pier (E) provided for reference.

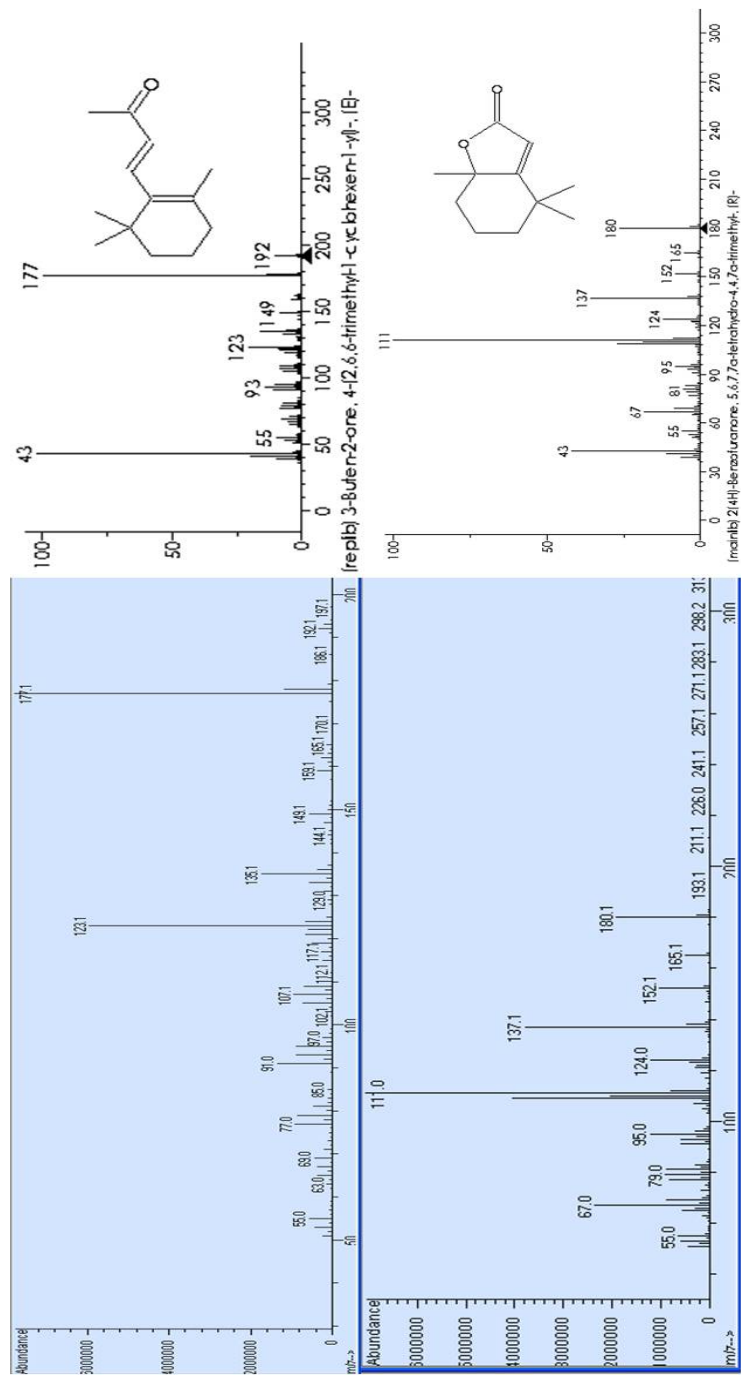


Figure 4.3: Observation of known  $\beta$ -carotene degradation products by GC-MS. Background subtracted mass spectrum from GC-MS on left side. Mass spectrum from NIST database on right side.

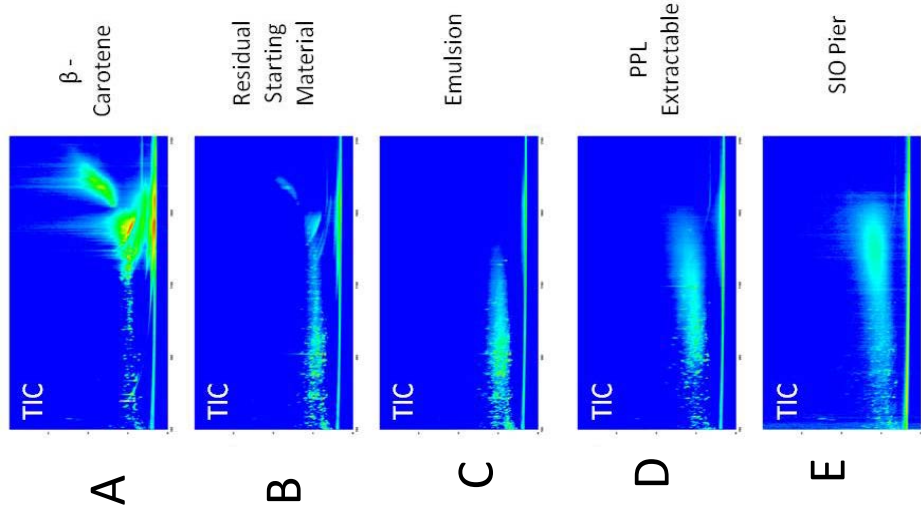


Figure 4.4. GCxGC-TOF-MS total ion chromatograms (TIC) of derivatized  $\beta$ -carotene (A), and oxidized  $\beta$ -carotene products Residual starting material (B) Emulsion (C) and PPL Extractable (D). Derivatized SIO Pier (E) provided for reference.

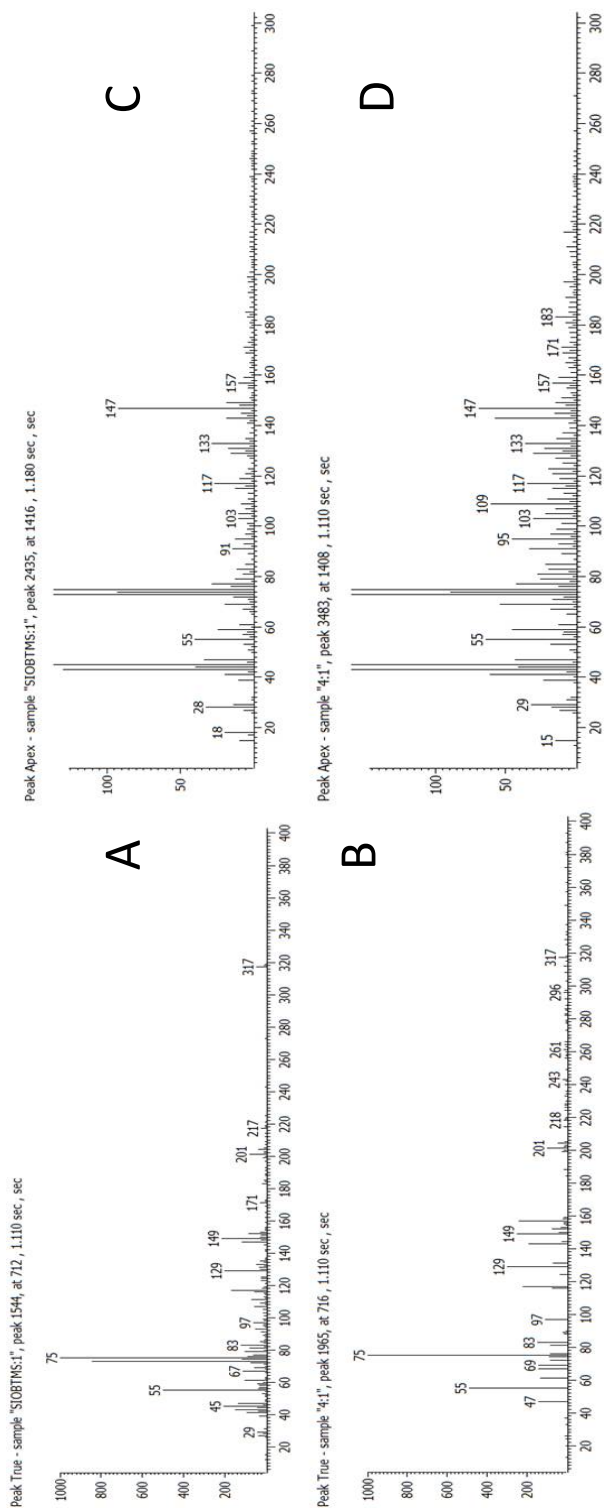


Figure 4.5: Mass spectra from PPL Extractable and SIO Pier. Peaks (A) and (B) are peak true (background subtracted) mass spectra of chromatographically similar peaks in SIO Pier and PPL Extractable. Peaks (C) and (D) are peak apex (spectrum at that time point without background subtraction), from a chromatographically unresolvable area. Both sets of spectra indicate similarity in observed compounds, but not necessarily exact matches.

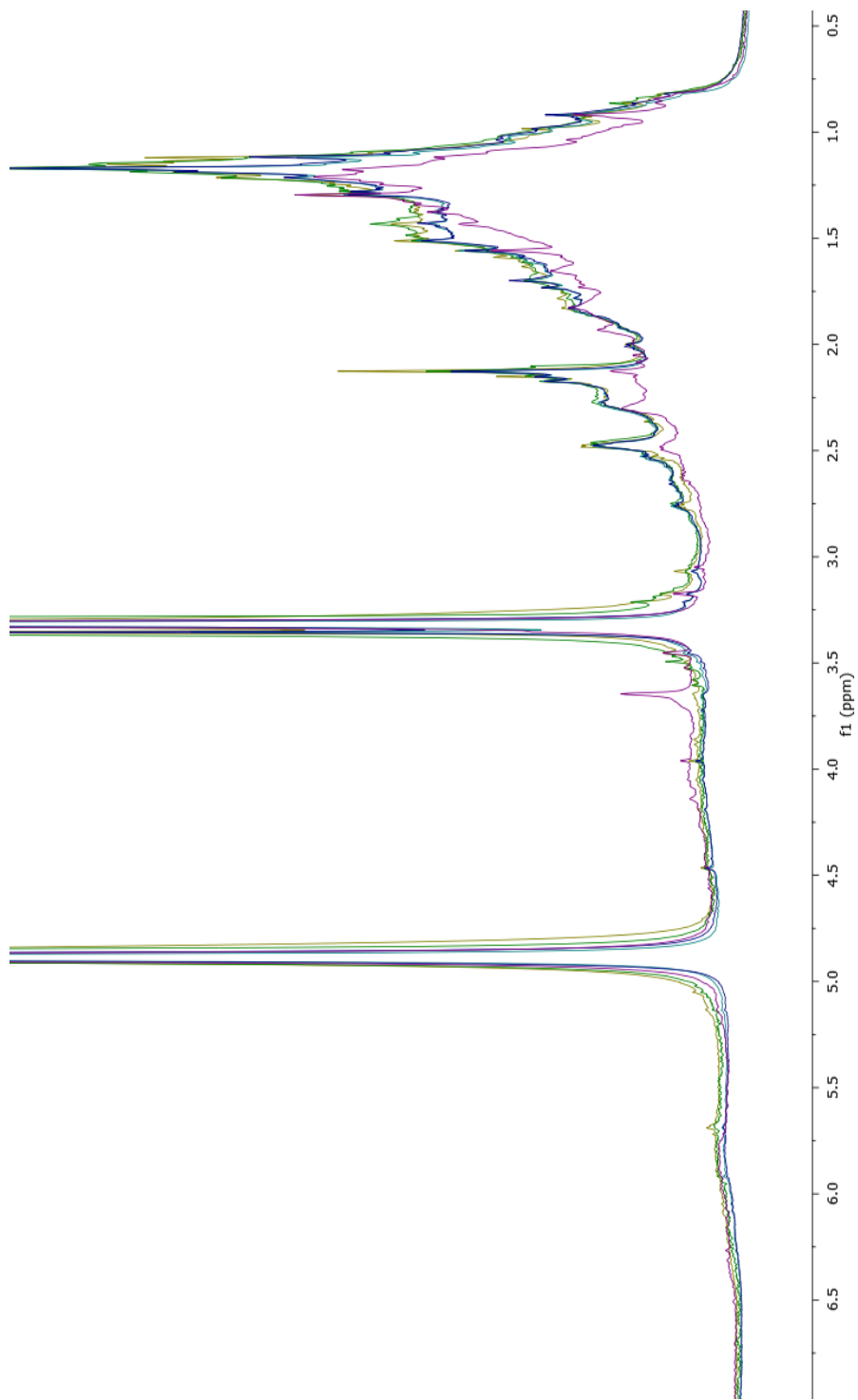


Figure 4.6.  $^1\text{H}$  NMR spectra of ‘Dark’ (42d) control, purple, and light PPL extractable fractions: 7d, green; 24d, yellow; 42d, light blue, dark blue.

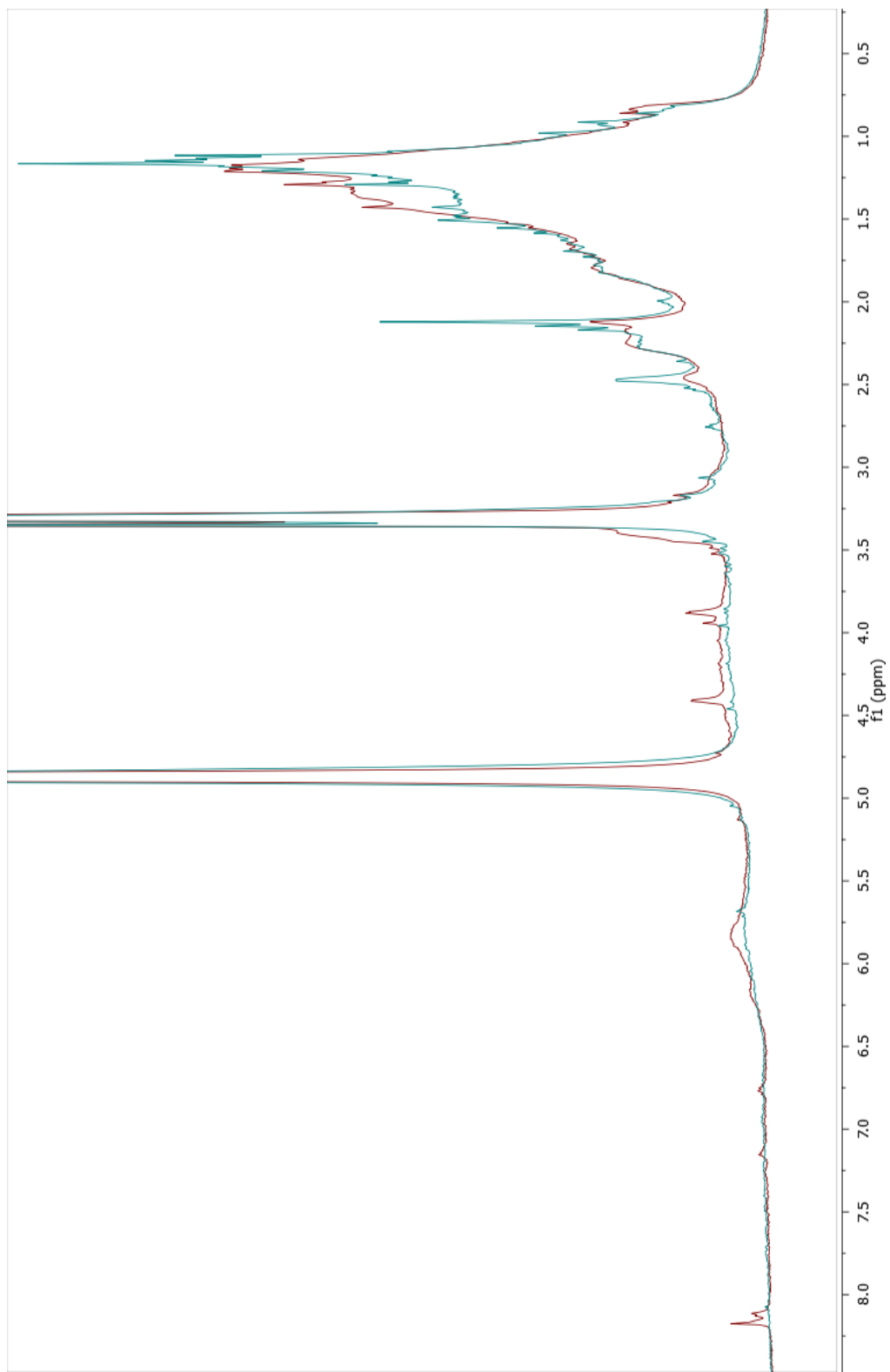


Figure 4.7: <sup>1</sup>H NMR of filtered Seawater PPL Extractable (blue) vs MilliQ PPL Extractable (Red)



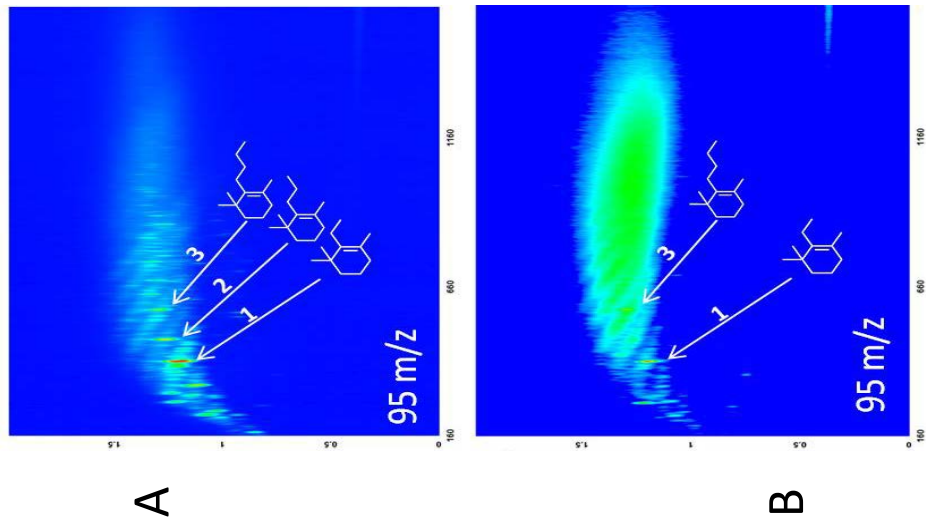


Figure 4.8. 95 m/z of reduction products of PPL extractable  $\beta$ -carotene degradation products, and SIO Pier. Peaks 1 and 3 represent compounds whose structure has been confirmed by model compound reduction. The structure of Peak 2 is strongly inferred based on retention times and mass spectrum (Fig. 4.9)

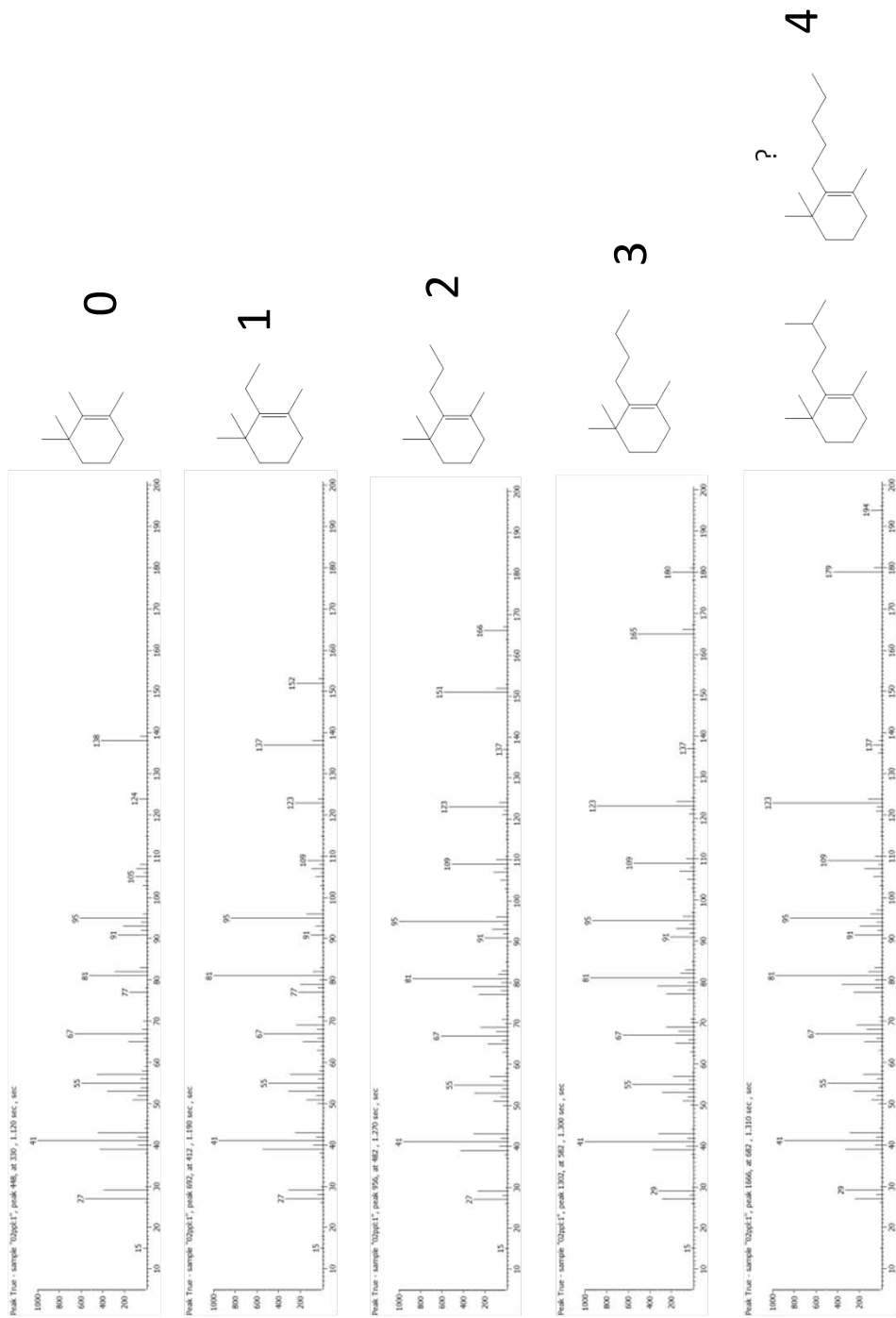


Figure 4.9. Mass spectrum of compounds identified in reduced PPL Extractable products. Peaks with these spectra are evenly spaced in GCxGC chromatogram (Fig 4.8), indicating that they are a series of related compounds. Structures on the right are either confirmed (1, 3), or hypothesized (0, 2, 4)

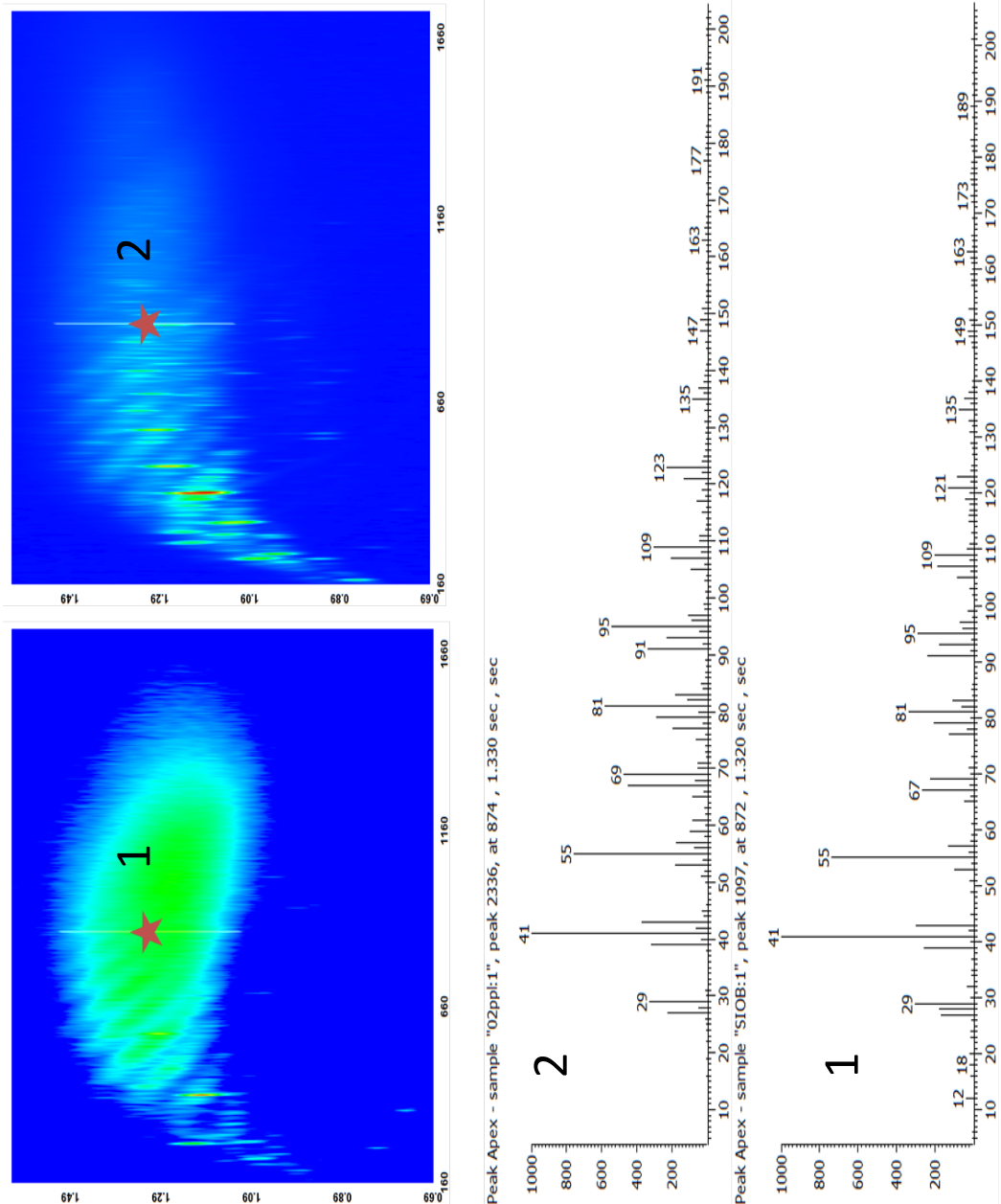


Figure 4.10. Peak apex (spectrum at point in chromatogram, not background subtracted) of unresolvable regions in SIO Pier (1) and PPL Extractable (2) with mass spectrum which demonstrate strong similarities in reduction compounds.

Table 4.1. EA-IRMS data of  $\beta$ -carotene (A), oxidized  $\beta$ -carotene products Residual starting material (B) Emulsion (C) and PPL Extractable (D) and SIO Pier (E) provided for reference.

	% Nitroge n	% Carbon	<sup>13</sup> C
$\beta$ -Carotene	0.01	90.62	-27.44
Residual Starting Material	0.01	73.36	-27.28
Emulsion	0.23	60.65	-28.33
PPL-extractable	0.11	57.00	-26.51
SIO Pier	2.12	48.09	-22.00

## References

- Dittmar, T. 2015. Reasons behind the long-term stability of dissolved organic matter in Biogeochemistry of Marine Dissolved Organic Matter 2<sup>nd</sup> edition. Hansell, D.A., Carlson, C.A., eds. Elsevier: 369-388
- Hansell, D.A., Carlson, C.A.. 2015. DOM Sources, Sinks, Reactivity, and Budgets in Biogeochemistry of Marine Dissolved Organic Matter 2<sup>nd</sup> edition. Hansell, D.A., Carlson, C.A., eds. Elsevier: 369-388
- Isoe, S., Hyeon, S.B., Sakan, T. 1969. Photo-oxygenation of carotenoids. I. The formation of dihydroactinidiolide and  $\beta$ -Ionone from  $\beta$ -Carotene. Tetrahedron Letters 4: 279-281
- Lam, M., Baer, A., Alae, M., Lefebvre, B., Moser, A., Williams, A., Simpson, A.J. 2007. Major structural components in freshwater dissolved organic matter. Environmental Science and Technology 41: 9240-9247
- Liaaen-Jensen. 1991. Marine carotenoids: recent progress. Pure & Applied Chemistry 63: 1-12
- Repeta, D.J., Gagosian, R.B. 1982. Carotenoid transformation in coastal marine waters. Nature 295: 551-54.
- Repeta, D.J., Gagosian, R.B. 1987. Carotenoid diagenesis in recent marine sediments-1. The Peru continental shelf (15°S, 75°W). Geochim Cosmochim Acta 51: 1001-1009

## **Chapter V**

Chemical characterization of solid-phase extracted dissolved organic matter across a salinity transect

## **Abstract**

We analyzed a suite of solid-phase extracted (SPE) dissolved organic matter (DOM) samples across a salinity transect to examine the transformation of DOM in these highly dynamic environments. Bulk elemental data (e.g. N/C ratio,  $^{13}\text{C}/^{12}\text{C}$ ) indicate mid-salinity inputs of nitrogen-rich compounds. GC-FID of derivatized samples indicates significant molecular-level differences between mid-salinity samples and high-salinity samples even though the compound classes present in each environment are very similar. This finding supports the fresh production of related biochemical classes in different salinity environments, but also suggests that conservative mixing does affect some of the underlying chemical composition observed across the gradient. Comprehensive gas chromatography mass spectrometry (GCxGC-TOF-MS) analysis of the same samples indicates that in particular, the relative contribution from small amino acids and sugar monomers increase across the salinity gradient. Production of fucose increased noticeably in mid-salinity samples, while relative concentrations of other known amino acid and sugar monomers stayed relatively constant. Chemical reduction identified carotenoid derived material in all samples but their relative molecular weight distribution changed discernibly toward bigger isomers in marine samples indicating that the degradation products of these compounds in the marine environments may be distinct from those in terrestrial environments, perhaps as a result of the major transformation mechanisms that differ among these environments.

## **Introduction**

Dissolved organic matter (DOM) is commonly isolated from aqueous environments using one of two means, ultrafiltration or solid-phase extraction (SPE). Ultrafiltration primarily extracts DOM by size, resulting in a fraction of DOM termed high molecular weight (HMW) DOM. This fraction has been examined using a variety of different techniques including nuclear magnetic resonance (NMR) spectroscopy ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^{31}\text{P}$ ), infrared spectroscopy, and electron microscopy for the bulk; and lipid, carbohydrate, amino acid, lignin, and benzene poly carboxylic acids following molecular level analyses of chemically degraded HMW DOM (Benner, 2002). Solid-phase extraction relies upon adsorption of compounds to nonpolar resins through hydrophobic interactions, and so typically recovers a more carbon-rich, smaller molecular weight fraction of the available DOM pool. Early studies employing the XAD and C-18 resins subjected this fraction of DOM to the same types of analyses listed above for HMW DOM (e.g., Gagosian & Stuermer, 1977, Druffel et al., 1992, Engbrodt and Kattner, 2005). A study employing DEAE cellulose resin subjected freshwater SPE DOM to sophisticated NMR analyses and concluded that linear terpenoid material had to be prevalent in DOM. (Lam et al., 2007) More recently PPL resins have been broadly applied and characterized primarily by fourier transform ion cyclotron mass spectrometry (e.g., Flerus et al. 2012, Stubbins et. al. 2010) (FTICRMS). Based on elemental ratios these studies have identified primarily highly reduced aliphatic compounds but a few peaks appear in regions consistent with carbohydrate and protein elemental ratios. Although FTICRMS has effectively demonstrated the complexity of DOM it has not successfully provided precise information on chemical structures. A few investigators have utilized NMR spectroscopy to examine the dominant composition of PPL extracted



DOM and have confirmed the importance of cyclic lipids that are functionalized with oxygen containing groups (Hertkorn et al. 2013). The only studies that have chemically manipulated PPL DOM for structural characterization have focused on lignin oxidation products (Medeiros et. al. 2016)

In this study we have analyzed a suite of PPL DOM samples isolated across a salinity transect. To observe and identify change across samples we used a combination of bulk analyses and molecular-level characterization following chemical degradation of the sample. Our results are presented with an initial focus on bulk analyses. Elemental analysis shows trends that are consistent with two end member mixing between marine and terrestrial DOM, but also indicates production in mid-salinity (Mid) samples. This departure from a conservative mixing model is further supported by gas chromatography (GC) flame ionization detection (FID) of derivatized samples, which observed increases in small molecule contributions to Mid samples. In the second section, we evaluate the use of comprehensive gas chromatography time-of-flight mass spectrometry (GCxGC-TOF-MS) as a tool to investigate the “small molecule” region, finding increased support for compositionally different DOC pools across the salinity gradient. In the third section, we focus on the utility of single compound quantification within a DOM sample, highlighting the production of the deoxysugar fucose in Low to Mid salinity samples. Finally, in the last section we apply a previously published chemical reduction to select samples to examine how alicyclic lipids, implicated in refractory DOM cycling, varied across this salinity gradient.

## **Materials and Methods**

### *DOM extraction and isolation*

Estuarine and open ocean samples were collected by solid phase extraction in May 2014 on the R/V Savannah by collaborators from the Skidaway Institute of Oceanography. At the same time Altamaha River samples (A#) were collected by small boat. To isolate DOM all samples were acidified and passed through a Bond Elut PPL cartridge is now a widely used method (e.g. Hansman et. al. 2015, Sherwood et. al. 2015, Lechtenfeld et. al. 2014) based on an established protocol (Dittmar et. al. 2008). A largely similar extraction and elution procedure is outlined in Chapters 3, and 4. Fig. 5.1 shows the sampling region, with locations for samples Sapes, Altes, C1, C3, C5, C7, C11, and C12. Samples A1, A2, A3, A4, A5, and A6 were isolated from the boxed region in Figure 5.1. Table 5.1 lists the isolated samples, along with the salinity and total DOC concentration at each site. Extraction efficiencies were calculated by collaborators based on DOC recovered following elution of the PPL cartridge with methanol. PPL DOC is the extraction efficiency converted to a carbon concentration based on the total volume of water that was extracted via PPL.

### *Elemental and Stable Isotope ( $^{13}\text{C}$ and $^{15}\text{N}$ ) Analysis*

Elemental and stable isotope measurements were made at the Scripps Institution of Oceanography using standard elemental analyzer isotope ratio mass spectrometry (EA-IRMS). Elemental carbon and nitrogen content of each sample is reported as %C and %N. Carbon to nitrogen molar ratios are determined from measured elemental compositions. Stable isotopes of carbon and nitrogen are reported in standard  $\delta$  notation (‰, per mil) relative to Pee Dee Belemnite (PDB) or atmospheric nitrogen gas. DON

(Fig. 5.2A) are calculated from PPL-DOC (Table 5.1) and N/C ratio (Fig. 5.2B). DOC and DON mixing lines (Fig. 5.32) assume linear, conservative mixing between two endmembers.

Conservative mixing curves for N/C and  $\delta^{13}\text{C}$  have to be scaled to PPL DOC concentrations (Abdulla et. al. 2013), and follow the general form of Eq. (1)

$$I_M = \frac{(f_{A5} * I_{A5} * \text{DOC}_{A5} + f_{C12} * I_{C12} * \text{DOC}_{C12})}{\text{DOC}_M}$$

where  $I_M$  is the conservative value (either N/C ratio or  $\delta^{13}\text{C}$ ) at a mixing point M, where  $f_{A5}$ , and  $f_{C12}$  are the fractions of a terrestrial water (sample A6, Table 5.1) and marine water (sample C12, Table 5.1) as calculated from salinity,  $I_{A5}$  and  $I_{C12}$  are the conservative (endmember) values of A5 and C12,  $\text{DOC}_{A5}$  and  $\text{DOC}_{C12}$  are the respective PPL DOC concentrations at A5 and C12, and  $\text{DOC}_M$  is the conservative mixing estimate (Fig. 5.2A) for mixing point M. Generating an accurate mixing curve for  $\delta^{15}\text{N}$  requires a conversion to DON, such that the equation changed to Eq. (2).

$$\delta^{15}\text{N}_M = \frac{(f_{A5} * \delta^{15}\text{N}_{A5} * \text{DOC}_{A5} * (\text{N/C})_{A5} + f_{C12} * \delta^{15}\text{N}_{C12} * \text{DOC}_{C12} * (\text{N/C})_{C12})}{\text{DOC}_M * (\text{N/C})_M}$$

where  $(\text{N/C})_{A5}$  and  $(\text{N/C})_{C12}$  represent the N/C molar ratios at A5 and C12, and  $(\text{N/C})_M$  represents the calculated N/C molar ratio at mixing point M (as calculated for Fig. 5.3B).

### *Hydrolysis protocol*

Bulk samples were weighed (~1mg, in MeOH, dried overnight at 60°C) into combusted 400 µL GC Vial inserts on a balance accurate to ±0.05 mg. 300 µL of 2M HCL was added to each insert, which was placed into a GC Vial w/ screw top cap and heated overnight at 90°C. Vials were then transferred to a CentriVap bench top vacuum concentrator and concentrated under reduced pressure at 50°C to dryness. MilliQ deionized water was then added to remove any residual acid and samples were taken to dryness again. This process was repeated as necessary (typically 2x total) until the pH measured was approximately equal to 4. For GC analysis samples were taken to dryness prior to derivatization.

#### *Derivatization procedure, Bulk and Hydrolyzed samples*

Bulk samples were weighed (~1mg, in MeOH, dried overnight at 60°C) into combusted 400 µL GC Vial inserts on a balance accurate to ±0.05 mg. Hydrolyzed (see above) and non hydrolyzed bulk samples were derivatized in 400µL (3:1 BSTFA10%TMCS: Pyridine, Ball and Aluwihare 2014). GC vials were capped and heated at 70°C for 1 hour. Samples were then analyzed by GC-MS/FID and GCxGC-TOF-MS

#### *GC-MS/FID Analysis*

An Agilent 7890A Gas Chromatograph system coupled simultaneously to an Agilent 5975C quadrupole mass spectrometer and a flame ionization detector was used for GC-MS/FID characterization of the sample. Splitless injection with 1 µL of analyte was used. Separation was performed on a 5% phenyl poly(dimethylsiloxane) column

(J&W 123-5731DB-ht, 30 m, 320  $\mu\text{m}$  i.d., 0.1  $\mu\text{m}$  film) with a temperature program from 100  $^{\circ}\text{C}$  ( hold time 2 min) to 200  $^{\circ}\text{C}$  (at 5  $^{\circ}\text{C}/\text{min}$ , hold time 0 min) to 320  $^{\circ}\text{C}$  (at 8  $^{\circ}\text{C}/\text{min}$ , hold time 8 min). Helium was the carrier gas with a constant flow of 1.8 mL/min. After separation, the effluent was split between the flame ionization detector (operating at 310 $^{\circ}\text{C}$ ) and the mass spectrometer (70eV ionization, scanning 50-750 m/z). Each FID signal was manually integrated in triplicate and the average value was used. The average standard deviation for triplicates was 0.6%. FID areas are either presented normalized to sample amount, or by percent FID values corresponding to each region of interest (where percent was calculated by dividing FID area of interest by total FID area).

#### *Comprehensive Two-Dimensional Gas Chromatography (GCxGC-TOF-MS)*

The GCxGC instrument used is a LECO Pegasus 4D GCxGC-Time-of-Flight Mass Spectrometer (GCxGC-TOF-MS). The term GCxGC refers to the use of two distinct columns in series that have different chemical selectivity. Compounds are separated primarily by volatility in the first column and polarity in the second. The analysis is comprehensive because all of the effluent from the first column is cryofocused and transferred onto the second column. Effluent from the second column is then analyzed by the TOF-MS, which benefits from high spectral acquisition speed (50 Hz-500 Hz). Finally, the data is compiled into a two-dimensional chromatogram that is visualized and processed by ChromaTOF<sup>®</sup> software. Both columns are housed within an Agilent 7890A Gas Chromatograph. The splitless inlet temperature is set at 300 $^{\circ}\text{C}$ . The first dimension column is a semi-polar Crossbond<sup>®</sup> diphenyl dimethyl polysiloxane column (Restek Rxi-17Si, 30m length, 0.2550mm i.d., 0.25 $\mu\text{m}$  film thickness). For

analysis of reduced products, the column was programmed to remain isothermal at 40°C for 1 min, and ramped to 315°C at 3°C per minute. The modulator temperature was offset by +30°C to the primary oven. The secondary oven (within the GC) housed the second dimension non polar Crossbond<sup>®</sup> dimethyl polysiloxane column (Restek Rxi-1, 1.58m length, 0.250mm i.d., 0.25µm film thickness). The secondary oven temperature was offset by +25°C to the primary oven. The modulation period was 2.5 s, with a hot pulse time of 1.05 s and a cool time of 0.2 s. For analysis of TMS-derivatized compounds, temperature program was as follows; 115°C for 3 min, 320°C at 8°C/min. Modulation period of 4s, 1.6s hot pulse time, 0.4s cool time. The carrier gas was Helium at a constant flow of 1.5 mL/min. The acquisition delay on the TOF-MS was set to 160s. The acquisition rate was set to 50Hz, with a range of 5-1000 m/z. Electron Ionization was run at 70eV.

#### *Amino acid and sugar standard derivatization*

A total of 24 amino acid and sugar monomer standards were selected to develop classification regions. Standards were individually prepared in MilliQ stock solutions, and then combined to form 4 different mixtures. Samples in each mixture were selected for a variety of amino acids, pentose and hexose sugars to be sure of identity. Mixtures were then transferred CentriVap bench top vacuum concentrator and concentrated under reduced pressure at 50°C to dryness, and derivatized. The following standards were used: D-galactose, *myo*-Inositol, D-glucose, D-xylose, D-fucose, D-glucuronic acid, D-glucosamine, L-Rhamnose, D-arabinose, glycine, β-alanine, D-phenylalanine, L-

phenylalanine, D-serine, L-threonine, isoleucine, L-leucine, L-alanine, L-glutamic acid, L-valine, D-tyrosine, L-aspartic acid, phthalic acid, mannopyranose.

### *Reduction procedure*

The reduction procedure directly follows Arakawa and Aluwihare (2015) which was modeled after Nimagadda et al. (2007). For reductions, compounds (~5 mg) were weighed to a flame-dried 2 mL vial equipped with a stir bar. Under Argon atmosphere, 10 mg of  $B(C_6F_5)_3$  was added (100  $\mu$ L of 100 mg  $B(C_6F_5)_3$ /1 mL dichloromethane solution). Immediately after addition, 100  $\mu$ L of n-BS was added, also under Argon. The mixture was then allowed to stir-overnight. Note that the sample is completely soluble in dichloromethane after the reduction. Following reduction the samples were treated to remove excess catalyst and siloxane by-products. Approximately 50  $\mu$ L of the reaction mixture was transferred into 250  $\mu$ L of conc. KOH/MeOH. After 1 minute the mixture was extracted with pentane (3x 100 $\mu$ L). The pentane fractions were collected, washed with  $H_2O$ , and dried over  $Na_2SO_4$ . The organic fraction was then dried under  $N_2$  to 100 $\mu$ L for GC-FID/MS and GCxGC-TOF-MS analysis.

## **Results and Discussion**

### *Bulk-level analysis of SPE-DOM*

All 14 samples from the study area were analyzed by elemental analyzer isotope ratio mass spectrometry (EA-IRMS), which quantifies carbon, nitrogen, and their stable isotopes ( $^{12}C$ ,  $^{13}C$ ,  $^{14}N$  and  $^{15}N$ ) in a sample. To initially test the strength of a

conservative mixing model, the concentration of DOC that the sample represents (PPL DOC, as calculated from total available DOC and extraction efficiency, Table 5.1) was plotted against salinity for each sample (Fig. 5.2A, blue diamonds). It is observed that for all of the samples where salinity exceeds 5 PSU, the measured DOC has a positive deviation from a conservative mixing line (Fig. 5.3.A, blue line). DON values for each sample (Fig. 5.2.A, red triangles) were calculated based on elemental data (N/C ratio) and PPL DOC concentrations. Again, the data show strong positive deviations from the mixing line (Fig. 5.2A, red line), the strongest of which occur in the 20-30 PSU salinity range. This consistent positive offset in both PPL-DOC and PPL-DON is interpreted to show nitrogen rich organic matter production in mid-salinity waters.

Similar observed positive deviations in both  $\delta^{13}\text{C}$  (Fig.5.2C) and N/C elemental ratios (Fig.5.2B) in HMW DOM across a salinity transect have previously been attributed to production (Abdulla et. al. 2013). The carbon isotopic signature ( $\delta^{13}\text{C}$ ) is an often used as a proxy to differentiate terrestrial from marine samples. Terrestrial organic matter is typically more depleted (more negative  $\delta^{13}\text{C}$ ) than marine organic matter (e.g. Hedges et al. 1997, Opsahl and Benner 1997, Goldberg et. al 2015), a characteristic largely attributed to the more depleted  $\delta^{13}\text{C}$  signature of the inorganic C that is used by terrestrial primary producers. An increase in  $\delta^{13}\text{C}$  (Fig. 5.2C) for mid-salinity samples is therefore, consistent with aquatic carbon production in the mid salinity region. The N/C ratio of a sample is also a good tracer of DOM inputs where terrestrial DOM is carbon rich and thus typically has low N/C ratios (Abdulla et. al. 2013, Perdue and Koprivnjak, 2007). Positive deviations in N/C ratio (Fig. 5.2B), suggest production of non-terrestrial,



nitrogen-rich compounds in mid-salinity regions. Finally, the measured  $\delta^{15}\text{N}$  of samples, another source marker of DOM composition (Goldberg, et. al. 2015), appears to be relatively constant across the transect. This indicates that in this ecological system, nitrogen compounds in marine and terrestrial SPE-DOM have similar isotopic signatures.

To complement EA-IRMS data, Gas Chromatography-Mass Spectrometry/Flame Ionization Detection (GC-MS/FID) was used as a semi-quantitative method to compare SPE-DOM samples. Derivatization is a necessary step to access hydrophilic compounds by GC-MS/FID, and the BSTFA process was selected for its versatility and ease of use (Ball and Aluwihare 2014). In addition to analyzing native PPL-DOM extracts by GC-MS, these same samples were also hydrolyzed in 2M HCL, then derivatized and analyzed by GC-MS/FID. The hydrolysis was employed to mimic previous studies of HWM DOM and SPE-DOM, which sought to increase the molecular level information obtained for each sample. As the hydrolysis is mild, it would be expected to cleave primarily ester bonds, potentially liberating more GC-accessible compounds with minimal chemical alteration. In GC-MS/FID chromatograms for all samples (bulk and hydrolyzed), we observed discrete peaks eluting up until ~20 minutes. We have termed this region the “small molecule” region of the chromatogram (Fig. 5.3). The composition of this region will be addressed in the following sections. After this region (20-40 minutes) is a hump of chromatographically unresolvable material, which can be termed an “unresolved complex mixture,” or “UCM.” For the remaining discussion it is assumed that all PPL DOM is accessible to GC-FID detection as either “small molecules” or “UCM.” This

assumption is supported by the complete dissolution of sample material in the derivatizing solution (BSTFA and pyridine).

To semi-quantitatively discern differences between samples the two regions – “small molecule” and “UCM” - were integrated for each GC-FID chromatogram. Flame Ionization Detection (FID) is the most quantitative method available for analyzing carbon-containing compounds by GC. However, as we lack true standards for most compounds, the data is best applied for relative (between samples), rather than absolute, quantitation. In Fig. 5.4A, the small molecule and UCM regions for bulk and hydrolyzed samples are expressed normalized to the sample amount derivatized (units: FID area/ mg sample). To compare compositional characteristics across salinity gradients the samples were binned into salinity “groups”. DOM samples isolated from waters with salinity less than 10 (n=4) were grouped into “Low” salinity samples; samples from waters where salinity was between 10 and 30 (n=3) were grouped into “Mid;” samples with salinity greater than 30 (n=6) were grouped into “High.”

Samples grouped in this manner, exhibited two trends. First, small molecule abundances increased from Low to Mid to High salinity samples. Second, for every sample, hydrolysis increased the area of the small molecule region as quantified by FID. The post hydrolysis increase in small molecules is interpreted to result from the redistribution of molecules from the UCM into the small molecule region. The average FID area under the UCM for each salinity ‘bin’ does not change significantly before and after hydrolysis (Fig. 5.4A) in part because of the large variability in UCM area between samples. Also, the FID response associated with the UCM is large compared to the

change in the small molecule region pre- and post-hydrolysis. Therefore, it is unlikely that the GC-FID signal would be sensitive to the redistribution of compounds from the UCM to the small molecule region. The overall increase in the small molecule region with salinity is consistent with the positive deviation from the expected two endmember mixing line for both PPL-DOC and  $\delta^{13}\text{C}$  signature across the salinity gradient. Together the data confirm the addition of new DOM in mid-salinity and high salinity regions.

In contrast, when comparing between salinity bins the UCM region of the chromatogram exhibited significant differences in FID area only between Mid and High salinity samples (Fig. 5.4A). Although the *fraction* of the FID represented by small molecule groups did not differ significantly between Mid and High salinity samples (Fig. 5.4B), the total FID response per mg of sample was greater for the High salinity samples. There are two possible explanations for this. The first is that more of the High salinity sample is amenable to GC analysis (more volatile) – this finding would also challenge the earlier assumption that all of the PPL-DOM is amenable to GC-FID. The second is that the UCM in particular, of High salinity samples, is compositionally distinct, such that there are more functional groups that can be derivatized. This is distinct from the first explanation because a single TMS group has a significant FID response. So the more TMS groups that are added to a sample the greater the FID response, even if the sample carbon is invariable. In either case, Mid and High salinity samples clearly differ in their composition.

Taken together, the data indicate that a simple two endmember mixing model cannot effectively explain the observed elemental and chemical composition of Mid

salinity samples. Specifically, elemental data indicate that Mid samples are isotopically distinct (Fig. 5.2C) from Low samples, and compositionally different (Fig. 5.4A) from High samples. One likely explanation is that the composition of all three bins is distinct. This data is further supported by the incorporation of our two endmember standards, Suwannee River Fulvic Acid (SRFA) and SIO Pier DOM (SIO-SPE). As shown in Fig 5.4A, both hydrolyzed samples have FID integrations similar to samples within the appropriate salinity bin but are clearly distinct from one another. Discernible compositional differences across a salinity gradient are consistent with Abdullah et al., (2013) who observed a sharp transition in composition at salinities  $>20$  in a setting similar to the current study.

#### *Small molecule region mass spectrometric analysis*

Samples were analyzed by comprehensive gas chromatography (GCxGC), coupled to time-of-flight mass spectrometry (TOF-MS) to further separate individual compounds and assist with structural characterization. Through the use of a second column, GCxGC-TOF-MS provides enhanced chromatographic resolution. Single ion-chromatograms are visualized as two-dimensional landscapes, in which individual peaks are identifiable as ellipses (Fig. 5.5A). The 73 m/z ion was selected for visualization and quantification, as it is a prominent fragment of Trimethylsilyl-derivatized (BSTFA) compounds. As shown in Figure 5.5A, compounds from the small molecule region elute over a 2-D space and are separated more effectively, which enables accurate identification. The UCM region is still irresolvable, but present now as a blob instead of

a broad hump. A suite of amino acid and sugar monomer standards (n=20, see experimental section for list) were also derivatized and analyzed by GCxGC-TOF-MS to confirm the identity of a subset of compounds common to each sample. In addition, these standards served to establish regions within the GCxGC landscape where related compounds, for which exact standards were unavailable, would elute. In total, 6 regions were established that encompassed most of the chromatographically resolvable small compounds.

The Pentasilylsugars region (2, Fig. 5.5A) was established based on the elution times of sugar monomer standards with 5 free alcohol groups, which incorporated 5 TMS-groups following derivatization. This group included mannose, galactose, and glucose. Two peaks corresponding to either the alpha or beta configuration at the anomeric carbon were detected for each sugar monomer. Both anomers were combined when quantifying each monomer. The Persilylsugars region (3, Fig. 5.5A) was established by sugar monomers with 4-TMS groups. These sugars include deoxysugars like rhamnose and fucose, and pentoses such as arabinose and xylose. Two other regions were created which contained compounds with mass spectra similar to TMS-derivatized sugars. The larger of these two regions has been designated the Trisilylsugar region (4, Fig. 5.4A). The smaller region has been termed the Silylsugar region (5, Fig. 5.5A). In a later section, we briefly present spectral and chromatographic evidence to support the identification of compounds within these two regions as methyl sugars. The two remaining regions were created based on the elution geography of amino acid standards. The first region (1, Fig. 5.5A) contained all of the amino acid standards (e.g., glycine, D-

serine, isoleucine). Typically, amino acids are derivatized with 3TMS-groups, with 2-TMS groups on the amine group and one on the carboxylic acid. Thus we speculate that compounds within this chromatographic region are generally derivatized with 3TMS-groups. The remaining group, Small Acids (6, Fig. 5.5A), encompassed compounds derivatized with 2 TMS-groups, although no standards were used to confirm the general structural characteristics of compounds eluting in this region. However, mass spectra are consistent with the general structure of small acids.

The ChromaTOF software was used to first identify compound peaks within each classification region in order to compare characteristics of the 6 classification areas across the sample set. Subsequently, compounds were quantified on the basis of the abundance of the 73 m/z ion that was present in the mass spectrum of each compound that was derivatized with BSTFA. The total area of all peaks integrated by using the 73 m/z ion, normalized to mg. sample (Fig. 5.6) is roughly in agreement with estimates made by GC-FID (Fig. 5.4A). Low salinity samples have the lowest peak area and High salinity samples have the highest total peak area (Fig. 5.6). Surprisingly, using the abundance of the 73 m/z ion to examine differences between the bulk and hydrolyzed samples indicated that they are not statistically different. This is in contrast to the trends shown in Fig. 5.4A and Fig. 5.4B for FID quantitation. The FID responds to the carbon content of the sample and is thus directly proportional to the abundance of GC amenable compounds. The 73 m/z ion on the other hand is directly proportional to the number of derivatizable functional groups present in the sample. It is possible that this property of the sample does not change appreciably following hydrolysis.

Focusing on peak area (73 m/z) by different regions, we observed trends largely consistent with elemental and FID data. The Amino Acid region increased from Low salinity to High salinity samples, although bulk samples did have higher peak areas than hydrolyzed samples. This shift, however, is misleading; this region is dominated by aminocaproic acid, which is a likely contaminant that is released following hydrolysis and subsequently recognizable by GCxGC-TOF-MS. There is evidence for its pre-hydrolysis form in the Amino Acid region (in further discussion), but its post-hydrolysis form was excluded from the integration of the Amino Acid region because it was considered to be a contaminant. If this peak is included in the integration then Bulk and Hydrolyzed samples have similar abundances. An increase in the Amino Acid region from Low to Mid is consistent with positive deviations in PPL-DON (Fig. 5.3A) and N/C Ratio (5.3B).

A similar decrease from bulk to hydrolyzed samples is observed in the Small Acid region (Fig. 5.7). This is not unexpected given the nature of compounds in this region. We would expect that a portion of these compounds would evaporate under drying conditions post hydrolysis. Once again, an increase from the Low to Mid groups is consistent with production of new DOC in the Mid region as inferred from PPL-DOC concentrations (Fig. 5.2A), isotopic (Fig. 5.2C) and FID (Fig. 5.4) data

Increases from Low to Mid to High salinity are observed in the combined Sugar regions, consistent with trends observed for the Amino Acids and Small Acid groups. However, the increase from Mid to High was significantly higher than that observed for Amino Acid or Small Acid regions, which indicates that the carbohydrate content and

composition of marine DOM is distinct. The breakdown of all sugars by region indicated that abundance increases in all fractions over the salinity transect.

Together the data presented in this section identifies differences in the small compound region of Low to High samples indicated by elemental (Fig. 5.2) and FID data (Fig. 5.4A). Prior to hydrolysis the GC-FID indicated a total increase in small compounds (Fig. 5.4A) from Low to Mid that reflected the production of new DOM. The unique compositional fingerprint of each sample as detected by GCxGC-TOF-MS confirms the FID data and is consistent with the observed positive departure of N/C ratios and  $\delta^{13}\text{C}$  values from the two endmember mixing line indicative of new DOM addition in the Mid region (Fig. 5.2). In GCxGC-TOF-MS, the increase is evenly distributed among all classification regions, with compounds in the Amino Acid region specifically contributing to some of the increase in DON. In contrast, the FID increase from Mid to High samples (Fig. 5.4A) appears to be driven by large relative increases in the sugar compound regions, which are more clearly detectable in the 1D chromatogram. Once again, the existence of two distinct transition trends (From Low to Mid, and Mid to High) indicates three distinct DOC pools.

#### *GCxGC-TOF-MS Identification of standards and single compounds*

In this section we move to a more targeted analysis of the small molecule region than previous sections, focusing on individual compounds. Compounds whose retention times and mass spectra were consistent with the amino acid and sugar monomer standards that were analyzed could only be identified in hydrolyzed samples. The relative abundance of amino acids (Fig. 5.8A) and sugars (Fig. 5.8B) is largely consistent



across Low, Mid and High salinity samples. Amino acids show a generally consistent composition across Low, Mid, and High salinity regions. This is not unexpected given that even functionally different proteins have similar amino acid compositions (e.g., Aluwihare and Meador, 2008). The relatively constant composition of individual sugar monomers is only disrupted by the clear production of fucose in Mid salinity regions. An increase in fucose is consistent with previous studies that identified deoxysugar production by unique phytoplankton communities that develop in mid salinity regions (e.g., Abdullah et al. 2013, Meador and Aluwihare, 2014). Fucose production, based on both relative abundance and normalized peak area (Fig 5.8C), is identified at salinities as low as 3.7 and is sustained throughout.

The lack of identifiable sugar or amino acid standards in bulk samples demonstrates that these compounds were inaccessible prior to acid hydrolysis. If the sugar monomers were contained within hetero biomolecules such as glycolipids, we would expect to see abundant fatty acids, which are not present. If fatty acids were present then they would have to be quite volatile and were lost during the centrivap step used to dry samples prior to derivatization. It is also possible that a fraction of the compounds that become “identifiable” following hydrolysis were originally isolated onto SPE cartridges as methyl esters. For example, aminocaproic acid is a compound that has been previously confirmed by GCxGC-TOF-MS of hydrolyzed SIO Pier samples (Fig. 5.9A), and identified in hydrolyzed samples in high abundance. The seemingly random (Fig. 5.9B) distribution, and lack of ecological significance, indicates that the compound is a contaminant. Like sugar and amino acid standards, aminocaproic acid is only

identified post hydrolysis. Pre-hydrolyzed samples contain an abundant peak with a similar mass spectrum (Fig. 5.9D) to derivatized aminocaproic acid (Fig. 5.9C). Given that the observed unknown peak in the bulk chromatogram has shorter 1<sup>st</sup> and 2<sup>nd</sup> dimension retention times, we are relatively confident that it represents the methyl ester version of aminocaproic acid. We dismissed the possibility that aminocaproic acid is modified by a methyl group on the amine based on the hypothesis that methylamines are far more hydrolytically stable than a methyl ester, and would thus be expected to survive the hydrolysis.

The abundance of peaks with sugar type mass spectra is an interesting feature of hydrolyzed SPE-DOM samples. It is best explained by the existence of several different methylated sugars in SPE-DOM. In Fig. 5.10 we have extracted three peaks from the Pentasilylsugar, Persilylsugar, and Trisilylsugar regions. The top peak corresponds to a TMS-derivatized mannose isomer. The middle peak corresponds to a TMS-derivatized xylose isomer. The last peak corresponds to an unknown compound. All three peaks have very similar mass spectra. The fact that xylose is a pentose, and has one less TMS-group than derivatized mannose, cannot be perceived from the mass spectrum alone. However, the increased size and the addition of a non-polar TMS group does extend the secondary retention time of mannose. Given that the unknown compound has a shorter secondary retention time than xylose, we can reasonably infer that it only possesses three TMS groups and therefore, has only three derivatizable groups. One shortcoming of electron impact ionization (for TOF-MS) is that parent ion information is lost. and so, the molecular weight of the compound is unknown. If we assume that the compound is a

sugar monomer (a reasonable assumption based on its mass spectrum), then the unknown compound is a deoxysugar with only three labile hydrogens (such as deoxyribose), a methylated sugar, or a methylated deoxysugar. We found that the Trisilylsugar region has the largest number of peaks of all the sugar regions, indicating a large diversity of compounds. The presence of ~50 methylated sugars and methylated deoxysugars (Panagiotopoulos et. al. 2007) has been identified in hydrolyzed HMW DOM, so it is reasonable to expect the same compounds to be present in hydrolyzed SPE-DOM. Given the mass spectral similarity of sugar derivatives, future work will require authentic standards to match chromatographic retention times and confirm the presence of methyl sugars.

GCxGC-TOF-MS analysis of standard amino acids identified similarities in the monomer composition of DOM across salinity regions but some informative differences also emerge. The relative distribution of compounds whose identity could be confirmed using ~20 standards did not change appreciably across the salinity gradient. Such compositional consistency may reflect a recalcitrant component of DOM that is inaccessible to biotic and abiotic degradation, and so, accumulates across these environments. Alternatively, and more likely, these monomers are common to different producers and environments. In contrast, fucose represents a single compound that tracks biological production associated with saline environments. Unfortunately, known compounds represent only a small fraction of the sample. All of the compounds in the abundant Trisilylsugar and Silylsugar regions, which we believe to be some combination of deoxy and methylated sugars, are presently unknown.

*Chemical reduction of selected samples*

Up to this point we have not discussed the composition of the UCM region of PPL DOM samples. In the previous three chapters the UCM was the explicit focus, and several lines of evidence suggested that a significant fraction of the UCM is derived from carotenoid degradation products. These structural features were identified in both terrestrial (Chapter 2) and marine (Chapter 3) samples. Carotenoids are abundant in all environments where primary producers are present. If carotenoid degradation products are part of refractory (unreactive) DOM then we may expect terrestrial carotenoids to be delivered to marine environments and subsequently accumulate there. The current sample set provided us with an opportunity to test this hypothesis and determine whether terrestrial and marine UCM compounds are similar or distinct.

A subset of samples was chemically reduced following established protocols. Two Low samples, two High samples, and a single Mid sample were reduced, along with SRFA and SIO-SPE (as end-member representatives). In previous reductions of SRFA and SIO-SPE we have observed similar alicyclic hydrocarbon products, which we have largely attributed to carotenoid degradation products. One difference between SRFA and SIO-SPE is the relative distribution of these products in 2D space. The intensity of red in the chromatogram of reduced SRFA (Fig. 5.11A), indicates where recovered products are most abundant. In the case of SRFA, the most abundant compounds elute earlier than ~900 seconds (white line). However, in reduced SIO-SPE (Fig. 5.11B), the most abundant material is centered around the white line. This indicates that SIO-SPE UCM contains carbon backbones that are larger (by intensity of distribution) than those found

in SRFA. This qualitative difference, mildly discernible in 2D chromatograms of the sample set (Fig 5.11 C-G), is much more apparent in 1D chromatograms of the sample set. In Fig 5.12 we observe that reduced SIO-SPE and SRFA have very different elution features, specifically in the 800-1400 second (box, Fig 5.12) region. Like the 2D chromatogram, an extended distribution of reduced SIO-SPE products indicates an intensity shift towards larger contiguous degraded carotenoid backbones. The reduction products of both samples A2 and A4, which are Low samples, are much more similar to reduced SRFA than reduced SIO-SPE in distribution. In contrast, the High samples, C3 and C7, are nearly identical in the retention time distribution to reduced SIO-SPE. Finally, the one Mid representative, SA, appears to be far closer in distribution to Low than High samples. If mixing were driving the composition of Mid samples, we would expect the distribution to look like a combination (and if so likely closer to High samples, given the salinity). As determined in previous chapters, quantitative reduction of SRFA and SIO-SPE indicate that the observed compounds represent at least 10-13% of the original sample. Although quantitative estimates were not made for this sample set, the products appear to be recovered in similar amounts. As such it should be stated that the distribution of reduced compounds does not represent the majority of unidentified compounds in the UCM. However, the distribution does indicate that the UCM of Mid samples may be more compositionally similar to Low than High samples, which is in agreement with GC-FID data (Fig. 5.4A)

## **Conclusion**

In our chemical characterization of solid phase-extracted dissolved organic matter across a salinity transect, we employed a combination of analyses that has not been previously applied to study SPE-DOM. These techniques range in analytical scale from bulk measurements to single compound analysis, yet the results are largely coherent. The sum total of the observations describes a sample set that is influenced by mixing, but is best characterized as three distinct pools of DOC. Mid salinity samples, expected to be a mixture of endmembers, is compositionally more similar to Low salinity samples (as evidenced by GC-FID, and reduced samples). However, Mid samples also contain elemental indicators of organic matter production, supported by the increase in the specific sugar fucose, which distinguish them from Low samples. Finally, peaks consistent with proteins and carbohydrates in chemical formula ratios have been identified in FTICRMS of SPE DOM, our results demonstrate that they may be an underappreciated component of SPE DOM. Single compound identification of putative methyl sugars is the most attractive route forward in characterizing other differences, and will rely upon standards and soft ionization techniques to confirm structurally unknown compounds.



Figure 5.1. Sampling region near Skidaway Institute of Oceanography, GA. Samples isolated primarily on orange transect. White box indicates additional area of sample isolation (samples Alt 1-6). Approximate scale: 0.75 inch = 10 miles

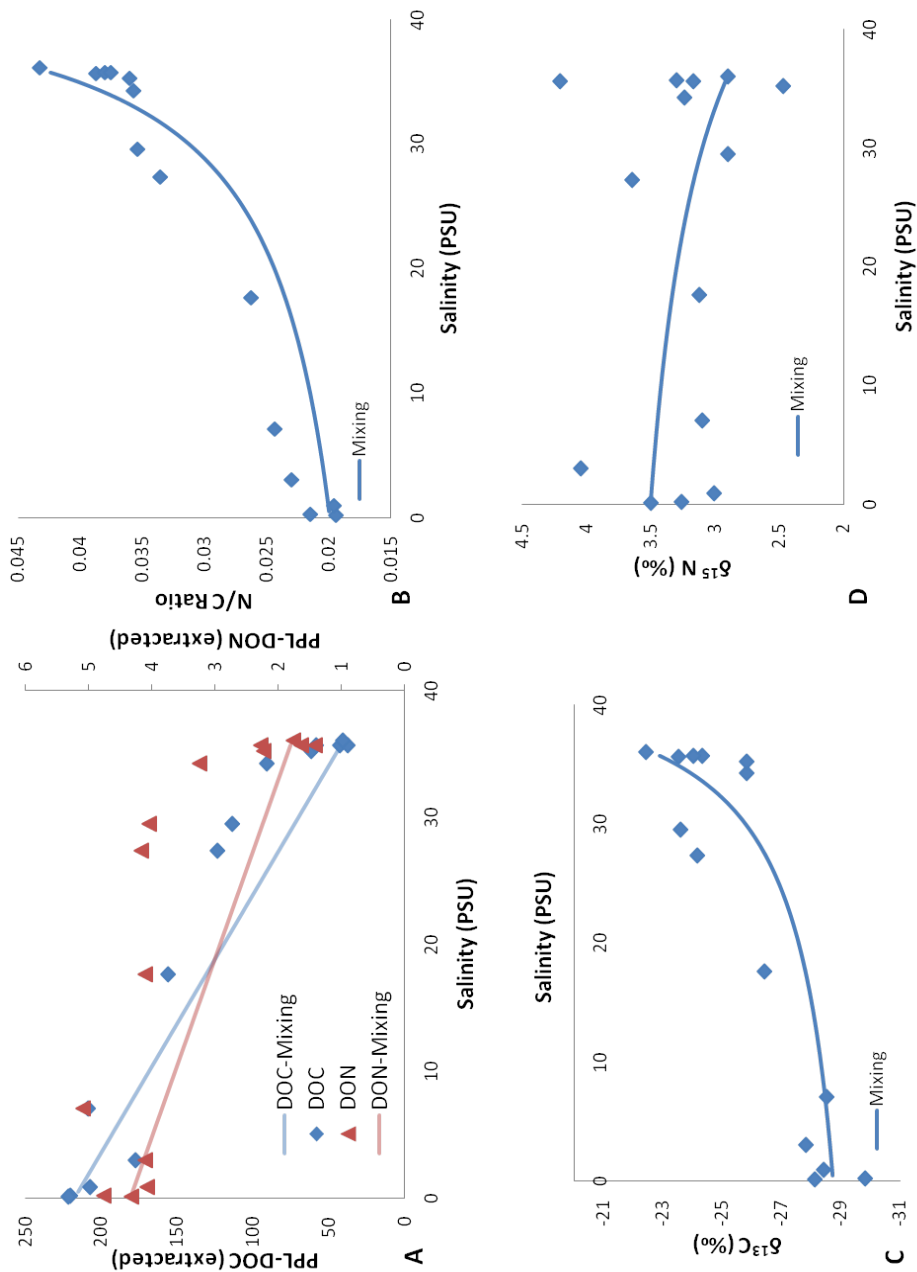


Figure 5.2. Elemental analyzer isotope ratio mass spectrometry (EA-IRMS) bulk data of sample set. Measured DOC (A, Blue diamonds), and calculated DON (A, red triangles) of extracted samples, and calculated conservative mixing lines. Elemental N/C ratios (B),  $\delta^{13}\text{C}$  (C),  $\delta^{15}\text{N}$  (D) for analyzed samples, and calculated mixing lines.



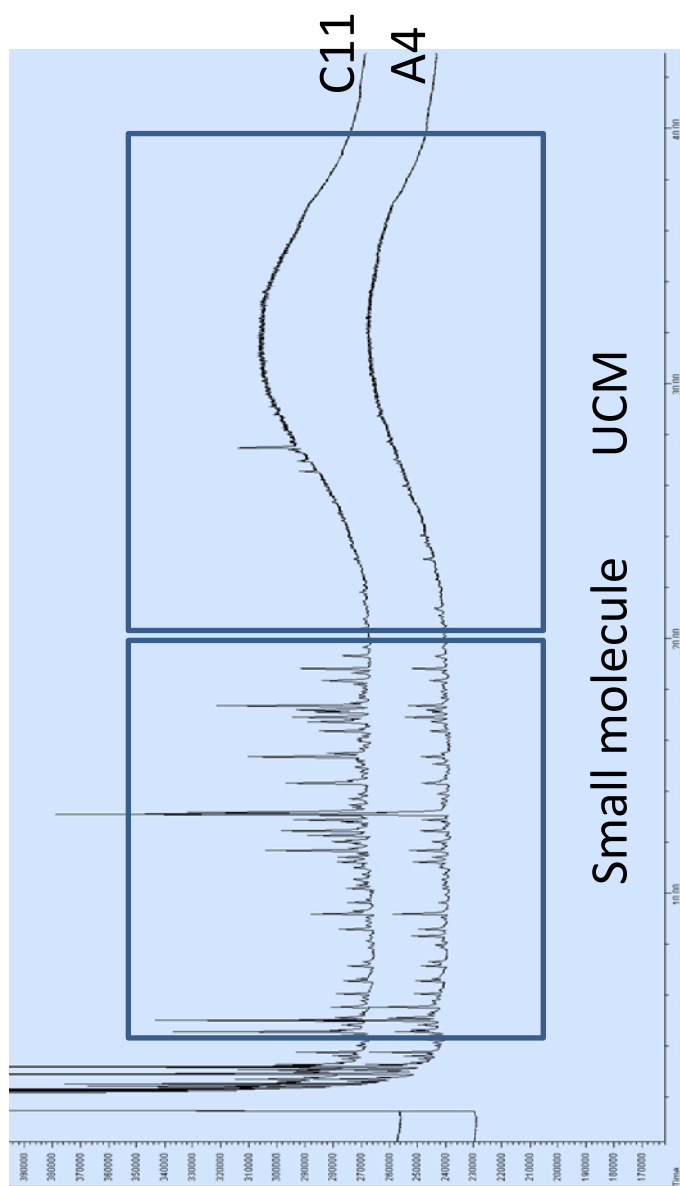


Figure. 5.3 Representative GC-FID chromatograms of sample set. C11 (top) is a high salinity sample, and A4 (bottom) is a low salinity sample. Boxed regions indicate elution windows for either “Small molecule” (0.5-20 min), or “UCM” (20-40 min) regions.

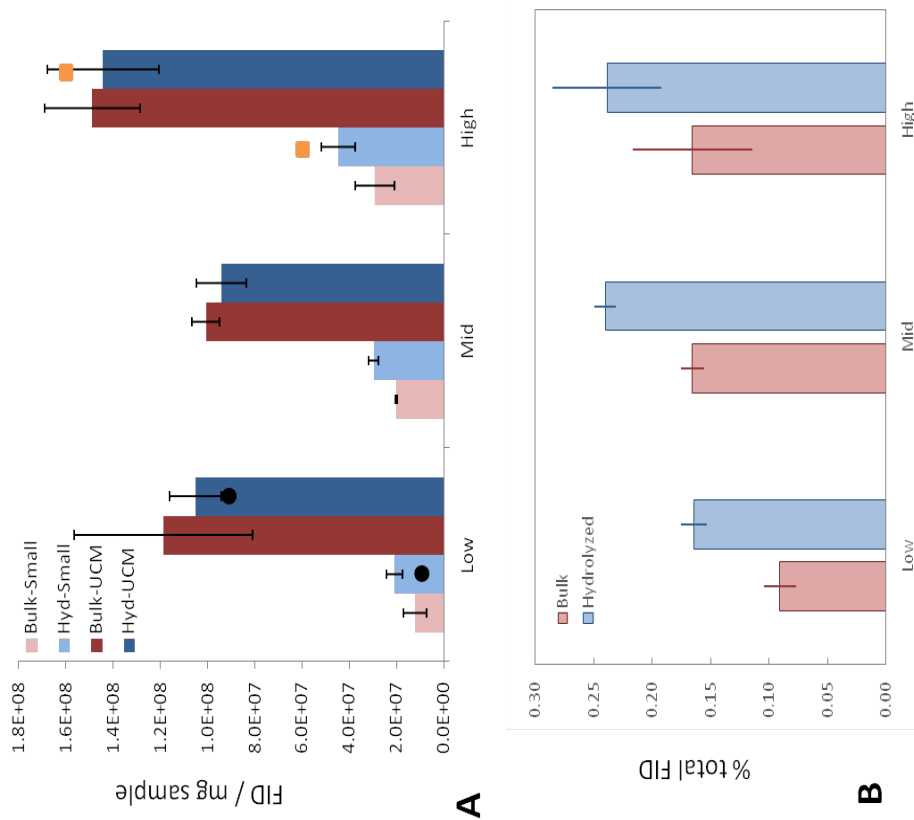


Figure 5.4. FID integrations of Small molecule and UCM regions for bulk and hydrolyzed samples, (A, top), binned into salinity regions : “Low,” 0-10; “Mid,” 10-30; “High,” >30. Columns indicate areas for bulk (red) and hydrolyzed (blue) sample sets, standard deviation ( $\delta$ ) indicated by error bars. Values for hydrolyzed SRFA (Black circles), and hydrolyzed SIO-SPE (Orange circles) also displayed. The small molecule region (B, top) is also presented as a fraction of the total (UCM+Small molecule).

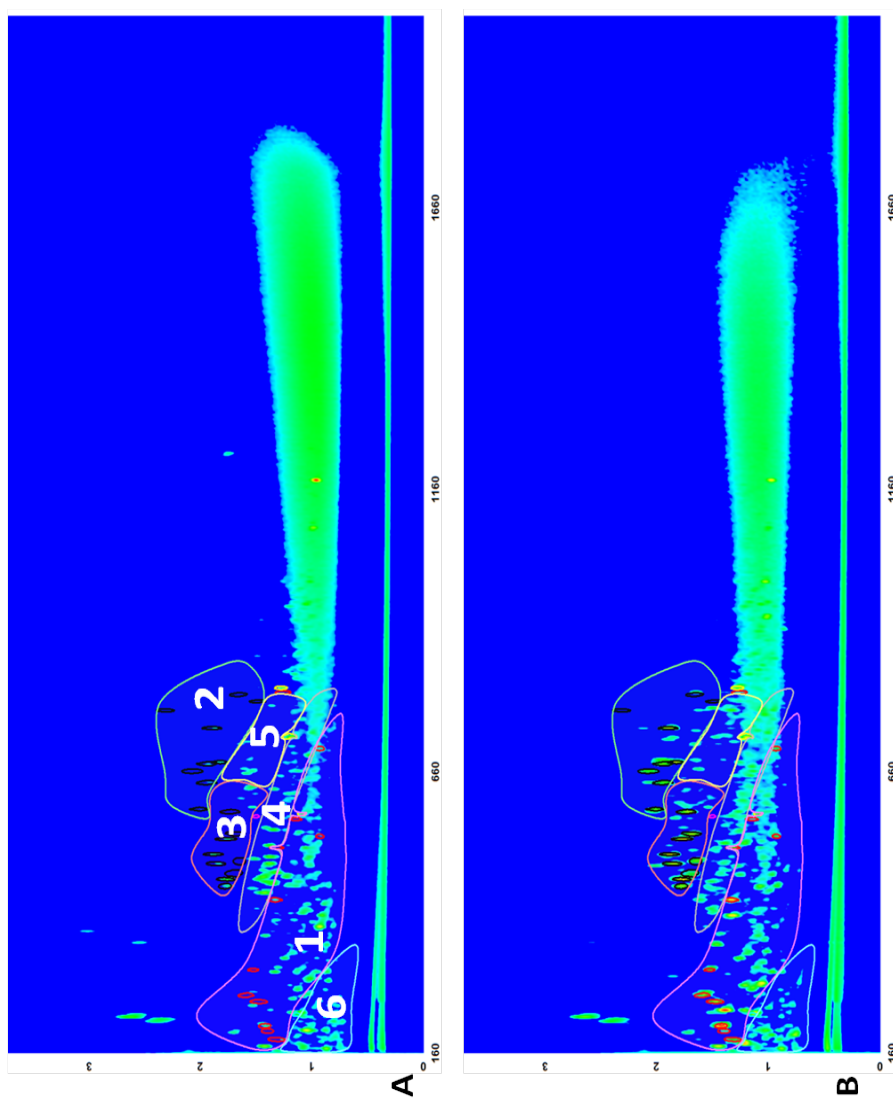


Figure 5.5. GC-TOF-MS 73 m/z chromatograms for A6 bulk derivatized sample (top, A), and A6 hydrolyzed derivatized sample (bottom, B). Regions are labeled as follows: 1. Amino acids. 2. Pentasilylsugars. 3. Persilylsugars. 4. Trisilylsugars. 5. Silylsugars. 6. Small Acids. Further description of regions in text. Black ovals represent sugar monomer standards, red ovals indicate amino acid monomer standards.

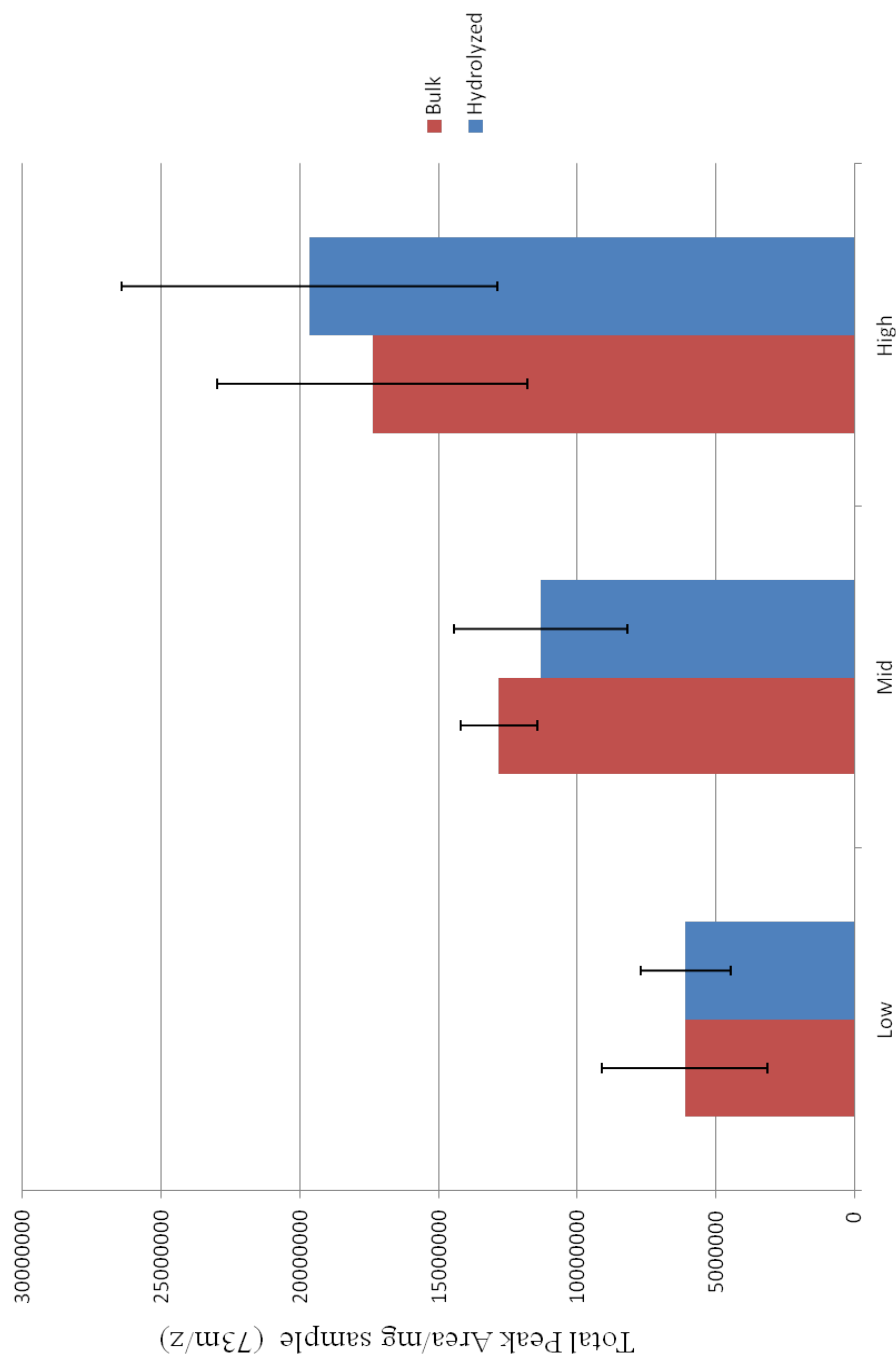


Figure 5.6. Total area of all peaks integrated with 73 m/z ion. Samples separated by bulk (red), and hydrolyzed (blue), and also binned into Low, Mid, and High regions. Error bars indicate standard deviation ( $\delta$ )

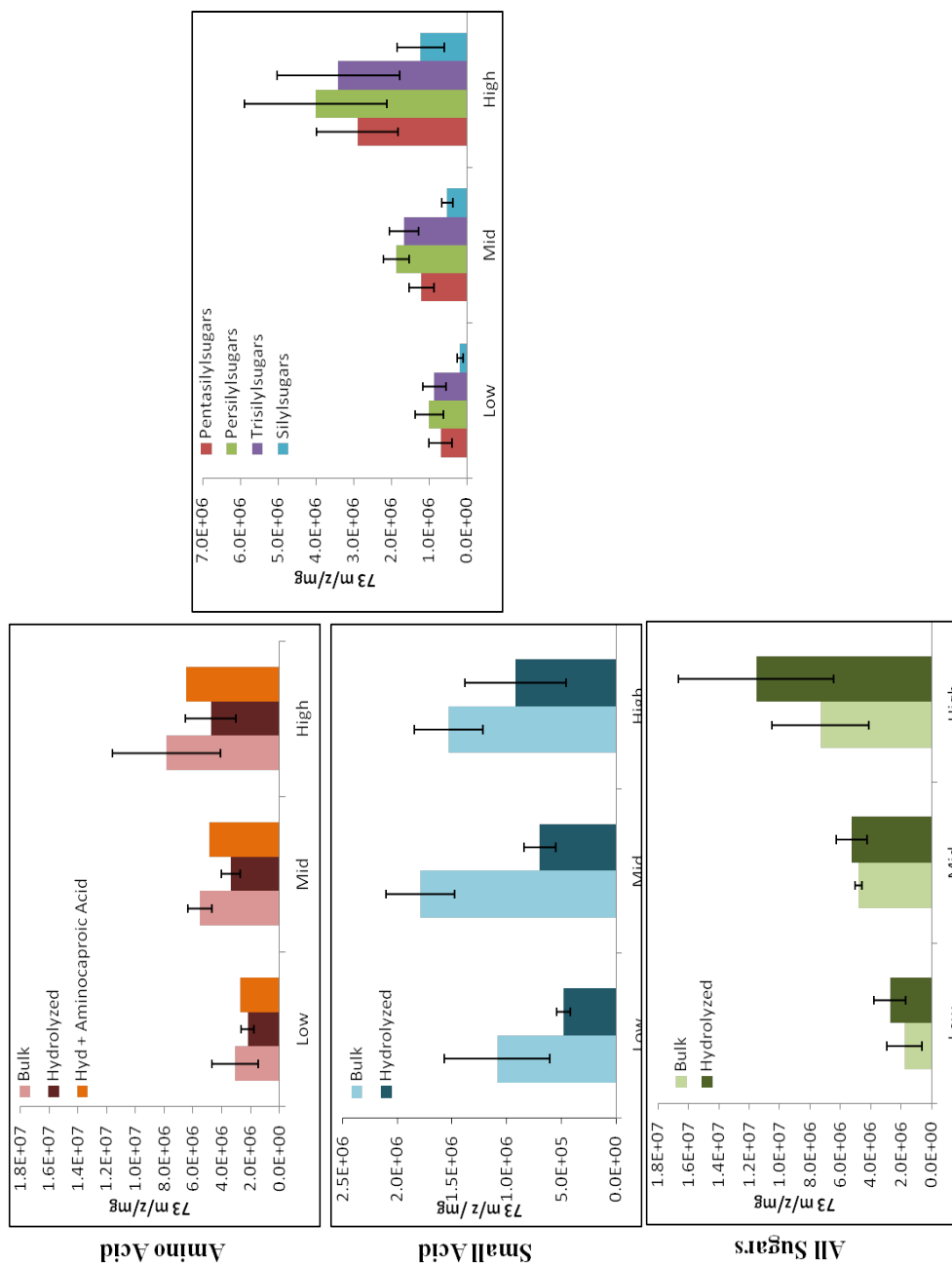


Figure 5.7. Total peak areas / mg sample as partitioned into classification regions. Samples were binned into Low, Mid and High groups. Break down of all sugars into different regions. Error bars indicate standard deviation ( $\delta$ )

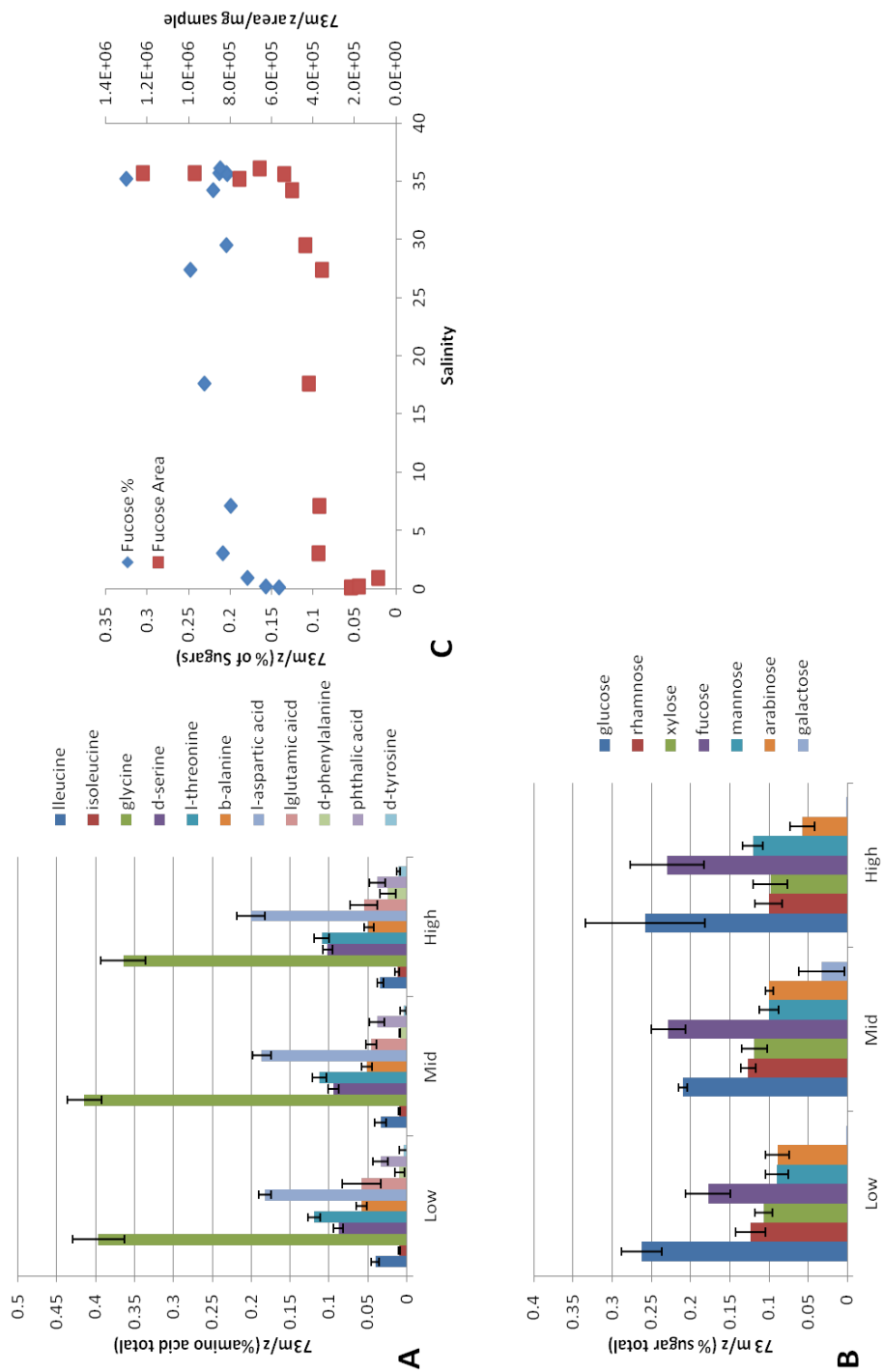


Figure 5.8. Total peak areas for samples as partitioned into classification regions. Samples were binned into Low, Mid and High groups. Fucose amounts given as % of total sugars, and area normalized to mg sample. Error bars indicate standard deviation ( $\delta$ )



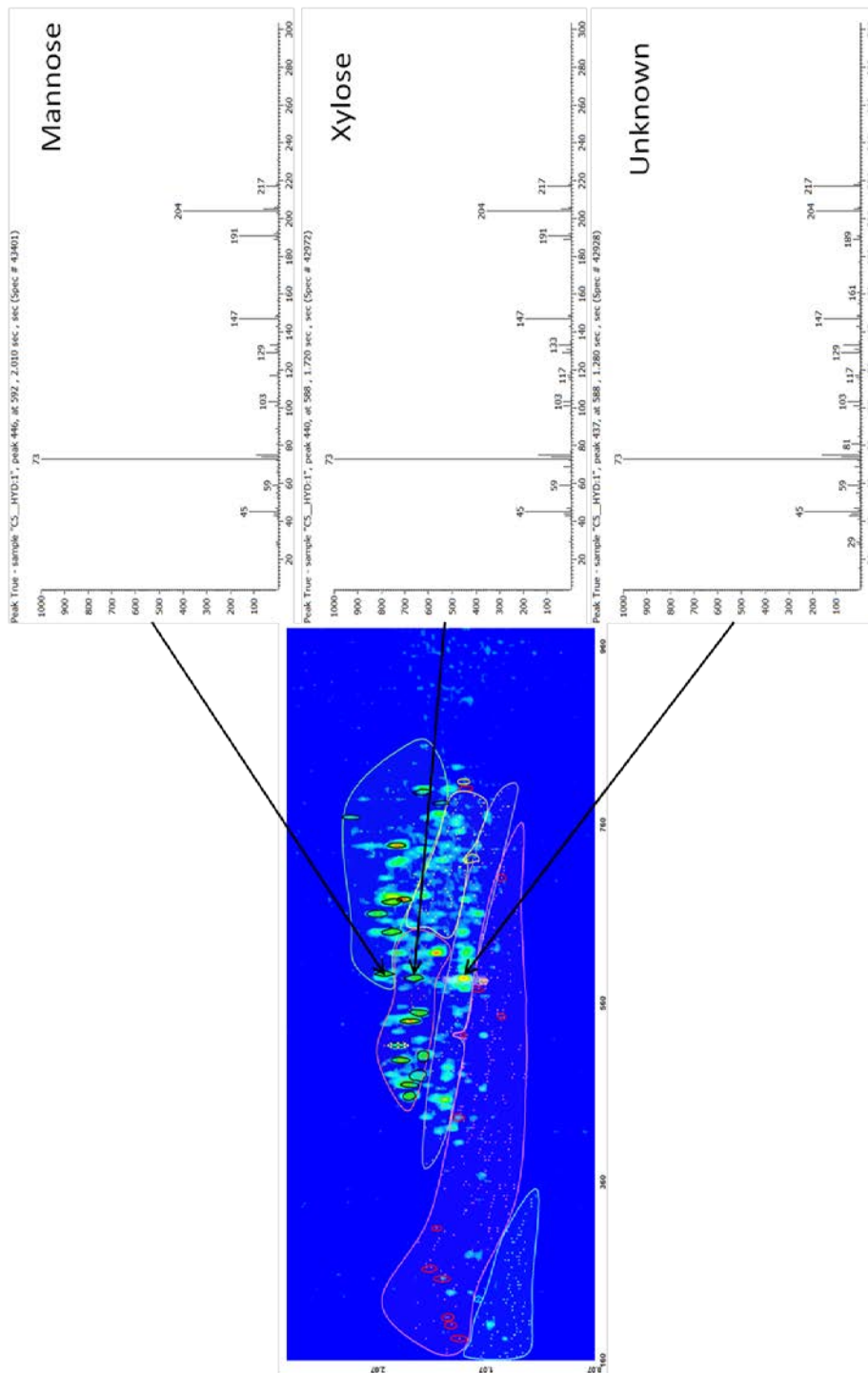


Figure 5.10. 204 m/z chromatogram of hydrolyzed C5 (High) sample, indicating sugar and sugar-like peaks in chromatogram. 4 regions with abundant peaks are Pentasilylsugars (top), Persilylsugars (left), Trisilylsugars (bottom) and Silysugars (right). Mass spectra for three peaks with different <sup>2</sup> RT times given at right. Mannose and Xylose are verified sugar standards.



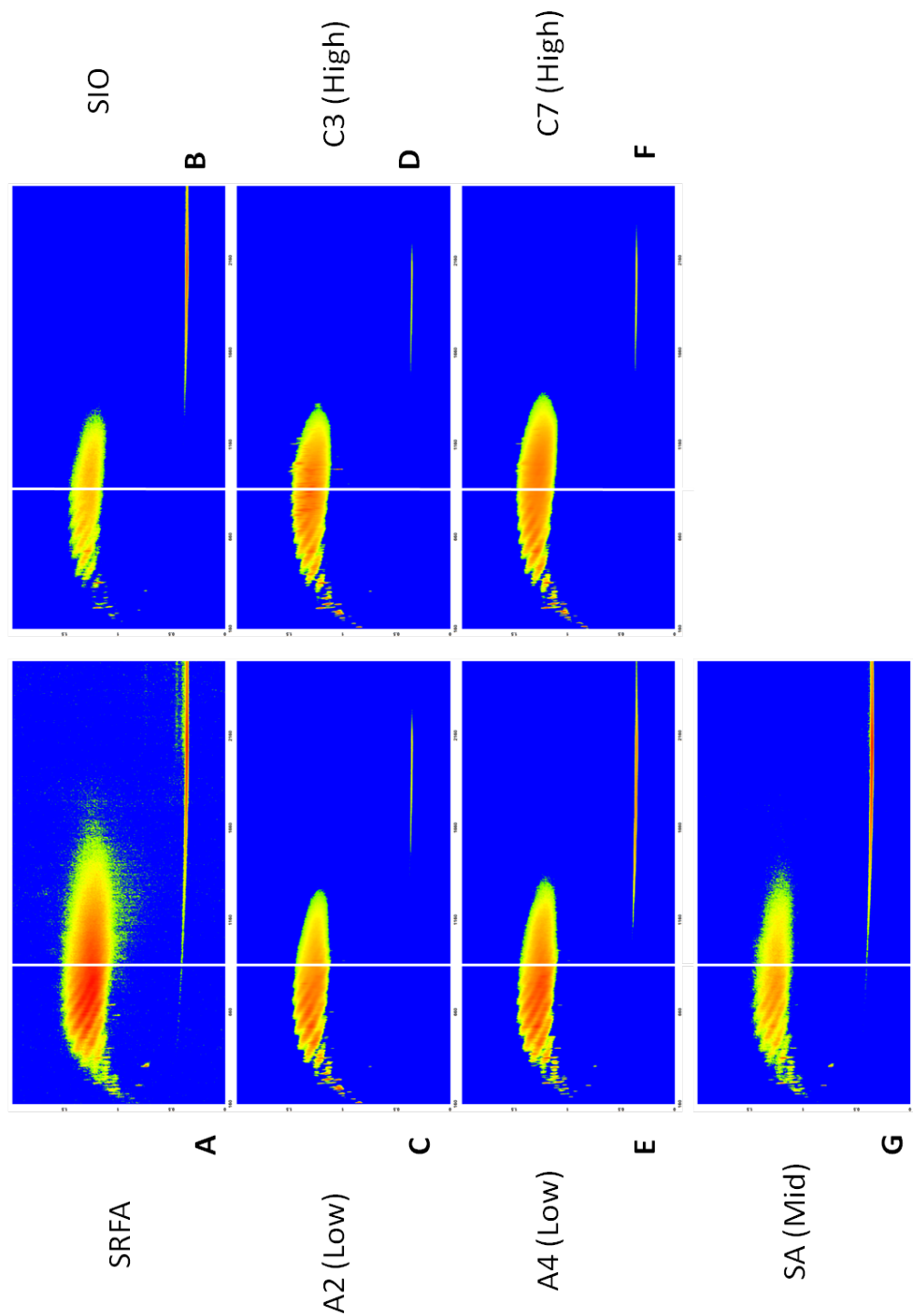


Figure 5.11. 95 m/z chromatograms of reduces SRFA (A), SIO-SPE (B), Low samples (C, E), High samples (D, F), and Mid sample (G). White line approximates  $900 \text{ s}^{-1}$  RT and is provided as a visual aid to distinguish elution patterns of reduction products.

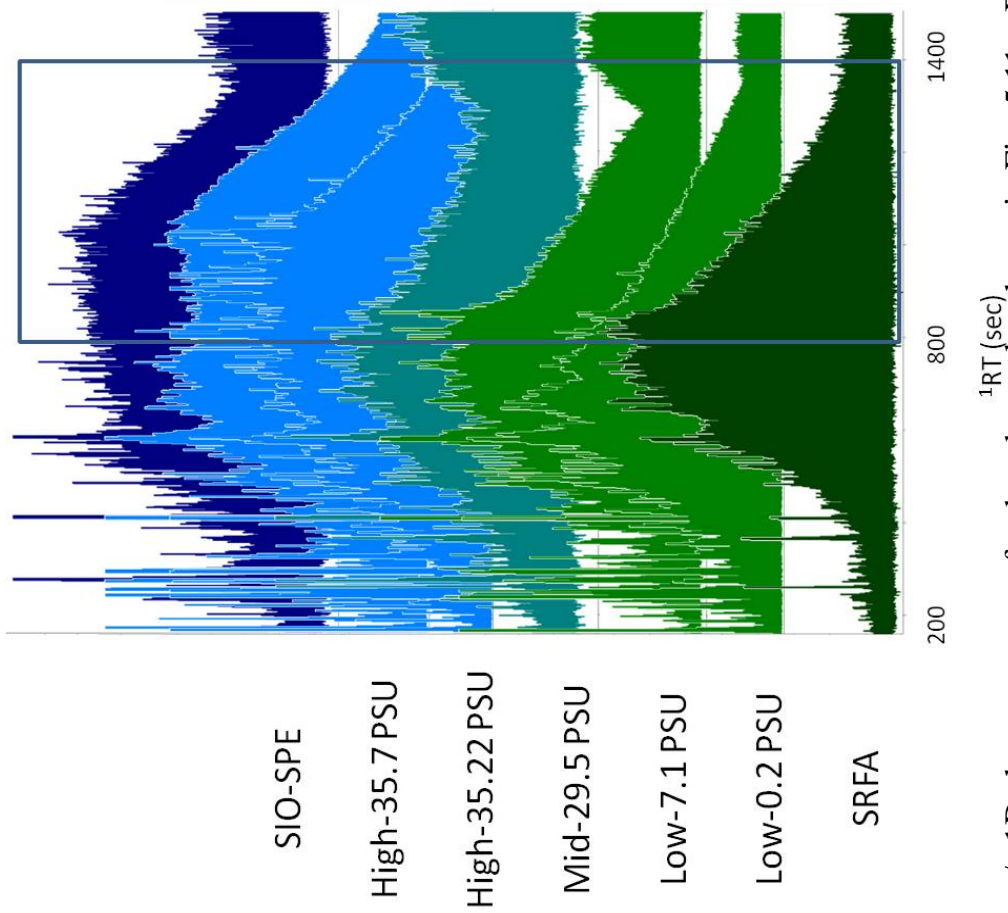


Figure 5.12  ${}^1\text{RT}$  95  $m/z$  1D chromatogram of reduced samples shown in Fig. 5.11 . Box from 800-1400 s  ${}^1\text{RT}$  to be used as visual aid in distinguishing sample set

Table 5.1. List of study samples. Sample [DOC] represents total DOC available. [PPL DOC] represents concentration of DOC that isolated sample represents ( $[PPL\ DOC] = [DOC] \times \% \text{ DOC}$ ).

Sample Name	Sample salinity	Sample [DOC] (uM)	% DOC extracted	[PPL DOC] (uM)
SAPES	29.5	286	40%	114
ALTES	27.3	212	59%	124
C1	34.3	166	55%	91
C3	35.2	114	55%	62
C5	35.6	122	49%	59
C7	35.7	95	46%	43
C11	35.7	88	44%	38
C12	36.1	86	47%	41
A1	17.6	264	59%	156
A2	7.1	366	57%	209
A3	3.0	379	47%	178
A4	0.2	368	60%	221
A5	0.1	358	62%	222
A6	0.9	362	57%	208

## References

- Abulla, H. A. N., Minor, E. C., Dias, R. F., Hatcher, P. G. 2013. Transformations of the chemical compositions of high molecular weight DOM along a salinity transect: Using two dimensional correlation spectroscopy and principal component analysis approaches. *Geochim. Comsochi. Acta* 118: 231-246
- Arakawa, N., Aluwihare. 2015. Direct Identification of diverse alicyclic terpenoids in Suwannee River Fulvic Acid. *Environmental Science and Technology* 49: 4097-4105.
- Ball, G.I., Aluwihare, L.I. 2014. CuO-oxidized dissolved organic matter (DOM) elucidated by comprehensive two dimensional gas chromatography-time of flight-mass spectrometry (GCxGC-TOF-MS). *Organic Geochemistry* 75:87-98
- Benner, R. 2002. Chemical composition and reactivity in *Biogeochemistry of Marine Dissolved Organic Matter*. Hansell, D.A., Carlson, C.A., eds. Academic Press: 59-90
- Dittmar, T., Koch, B., Hertkorn, N., Kattner, G. 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnology and Oceanography: Methods*: 6: 230-235
- Druffel, E.R.M., Williams, P.M., Bauer, J.E., Ertel, J.R., 1992. Cycling of dissolved and particulate organic matter in the open ocean. *Journal of Geophysical Research* 97: 15,639-15,659
- Engbrodt, R., Kattner, G. 2005. On the biogeochemistry of dissolved carbohydrates in the Greenland Sea (Arctic). *Organic Geochemistry* 36: 937-948
- Flerus, R., Lechtenfeld, O.J. Koch., B.P., McCallister, S.L., Schmitt-Kopplin, P., Benner, R., Kaier, K., Kattner, G. 2012. A molecular perspective on the ageing of marine dissolved organic matter. *Biogeosciences* 9: 1935-1955
- Gagosian, R.B., Stuermer, D.H. 1977. The cycling of biogenic compounds and ether diagenetically transformed products in seawater. *Marine Chemistry* 5: 605-632
- Goldberg, S. J., Ball, G. I., Allen, B. B., Schladow, S. G., Simpson, A. J., Masoom, H., Soong, R., Graven, H. D., Aluwihare, L. I. 2015. Refractory dissolved organic nitrogen accumulation in high-elevation lakes. *Nature Communicaitons* 6: 6347

- Hansman, R. I., Dittman, T., Herndl, G. J. 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
- Hedges, J. I., Keil, R. G., Benner, R. 1997. What happens to terrestrial organic matter in the ocean? *Organic Geochemistry* 27: 195-212
- Hertkorn, N., harir, M., Koch, B.P., Michalke, B., Schmitt-Kopplin, P. 2013. High-field NMR spectroscopy and FTICRMS mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosciences* 10: 1583-1624
- Lam, M., Baer, A., Alae, M., Lefebvre, B., Moser, A., Williams, A., Simpson, A.J. 2007. Major structural components in freshwater dissolved organic matter. *Environmental Science and Technology* 41: 9240-9247
- Lechtenfeld, O. J., Kattner, G., Flerus, R., McCallister, S. L., Schmitt-Kopplin, P., Koch, B. P. 2014. Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean. *Geochim. Cosmochim. Acta* 126: 321-327
- Meador, T. B., Aluwihare, L.I. 2014. Production of dissolved organic carbon enriched in deoxy sugars representing an additional sink for biological C drawdown in the Amazon River plume. *Global Biogeochemical Cycles* 38: 1149-1161.
- Medeiros, P.M., Seidel, M., Niggemann, J., Spencer, R.G.M., Hernes, P.J., Yager, P.L., Miller, W.L., Dittmar, T., Hansell, D. 2016. A novel molecular approach for tracing terrigenous dissolved organic matter into the deep ocean. *Global Biogeochemical Cycles* 30: 689-699
- Nimmagadda, R.D., McRae, C.R. 2007. Characterisation of the backbone structures of several fulvic acids using a novel selective chemical reduction method. *Organic Geochemistry* 38: 1061-1072.
- Opsahl, S., Benner R. 1997. Distribution and cycling of terrigenous dissolved organic matter in the ocean. *Nature* 386: 480-482
- Panagiotopoulos, C., Repeta, D. J., Johnson, C.J. 2007. Characterization of methyl sugars, 3-deoxysugars and methyl deoxysugars in marine high molecular weight dissolved organic matter. *Organic Geochemistry* 38: 884-896

- Perdue, E. M., Koprivnjak, J. 2007. Using the C/N ratio to estimate terrigenous inputs of organic matter to aquatic environments. *Estuarine, Coastal and Shelf Science* 73: 65-72.
- Sherwood, B. P., Shaffer, E.A., Reyes, K., Longnecker, K., Aluiwhare, L., Azam, F. 2015. Metabolic characterization of a model heterotrophic bacterium capable of significant chemical alteration of marine dissolved organic matter.
- Stubbins, A., Spencer, R.G.M., Chen, H., Hatcher, P.G., Mopper, K., Hernes, P.J., Mwamba, V.L., Mangangu, A.M., Wabakanghanzi, J.N., Six, J. 2010. Illuminated darkness: Molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry. *Limnology and Oceanography* 55: 1467-1477

## **Chapter VI**

### Conclusions

## Conclusions

Refractory DOM represents most of the marine DOC reservoir (Hansell, 2013) and is therefore a significant fraction of the global carbon budget. In spite of its size, the chemical origins of this reservoir are not known. High field nuclear magnetic resonance spectroscopy has been used to assign hypothetical structures to accumulating DOM (Hertkorn, 2006), however, while the identified features are informative, the hypothetical structures have no ties to abundant biochemicals. Thus we have little understanding of which biochemicals are accumulating in DOM, and what specific processes control the formation and ultimate removal of these compounds from the water column. The goal of this thesis was to identify structural features in marine refractory DOM that could be linked to abundant source biochemicals and also, to test, at the molecular level, the hypothesis, advanced by previous studies, that much of this reservoir consisted of highly oxidized alicyclic compounds. In this effort we have succeeded by demonstrating that degraded carotenoid compounds are prominent in DOM samples from different environments, including in refractory DOM. This thesis achieved its goal by using multiple analytical approaches – approaches that combined bulk analyses with chemical degradation techniques and high-resolution chromatography. Such approaches are just as important as developing new DOM isolation methods that can access a greater fraction of the accumulating reservoir. The major contributions and further research directions of each chapter are outlined below.

### *Chapter II*



Suwannee River Fulvic Acid (SRFA) is a readily available, widely studied environmental DOM standard, and was thus a good proxy for aquatic refractory DOM. In this chapter we demonstrated unequivocally that alicyclic hydrocarbon backbones are present in DOM, and that chromatographically isolated compounds represented 13% of the original carbon. Degraded terpenoids in SRFA had been postulated using carbon ( $^{13}\text{C}$ ) nuclear magnetic resonance spectroscopy (NMR), but specific chemical structures could not be supported (Leenheer and Rostad 2004). My study confirmed that terpenoid backbones do exist in SRFA as alicyclic hydrocarbons of varied contiguous carbon length. High-resolution mass spectrometry of SRFA (Stenson et. al. 2003) had also previously observed series of molecular ions with prominent one carbon offsets (as  $\text{CH}_2$ ,  $\text{CO}_2$  differences). Such offsets were observed at the molecular level as prominent features in my data set, indicating that the degradation of alicyclic terpenoids represents a previously unknown pathway to producing compositional diversity in DOM. The isomeric diversity of observed reduction products is a complicated feature to interpret. Model compound reductions indicated that isomers of a specific molecular formula could be produced during the reduction. However, the range of isomers observed in reduced SRFA was much larger than has been observed for model compounds, indicating that part of the isomerism is native to SRFA. Fragmentation studies of SRFA have indicated that structural diversity may not be high (Witt et. al. 2009), which is seemingly at odds with our observed data set. However, the smooth distribution of isomers in reduction products demonstrated only small structural differences that were unlikely to be observed in MS fragmentation studies. The reduction of bald-cypress tree litter resulted in prominent resin acids and sterols, both of which could conceivably be oxidized to form SRFA (a

hypothesis proposed in Leenheer et. al. 2003). Unfortunately, specific structures could not be determined for reduction products. By extending this analytical approach to other environments, specifically marine DOM, I hoped to provide more certain structural insights, and thus moved forward to Chapter III.

### *Chapter III*

Marine refractory DOM was isolated from the Scripps Institution of Oceanography (SIO) Pier. The radiocarbon age of the isolated DOM was ~1625 years old. While deep DOM would have been more representative of refractory DOM, its concentration is actually expected to be very similar throughout the water column. The drawback of working with surface DOM is that it contains a higher concentration of recently produced biochemicals, thus potentially increasing the complexity of DOM. However, previous studies had suggested that the solid phase extraction method that I was applying to isolate DOM consistently isolated a representative fraction of the refractory DOM even in surface waters. The chemical reduction of SIO DOM resulted in alicyclic hydrocarbons similar to those observed in Chapter II, and isolated compounds represented 10% of the original carbon. In addition to the previously observed unknown products, I identified two compounds for which exact structures were determined. These two compounds had a carbon backbone reminiscent of degraded carotenoids, and a single study from freshwater environments had mentioned that linear terpenoids, perhaps similar to carotenoids, were responsible for the aliphatic resonances identified in aquatic DOM (Lam et al., 2007). As such, I decided to pursue the possibility that carotenoids, an abundant and reactive biomolecule, were the source of refractory nuclear magnetic

resonance (NMR) resonances previously identified in marine (Hertkorn et al., 2006) and terrestrial (Lam et al., 2007) DOM. I confirmed their direct contribution using model compounds with the same carbon backbone, which when reduced, produced compounds identical in their retention time and mass spectrum to those accumulating in marine DOM. As such, I proposed that carotenoid degradation products are an important component of the alicyclic fraction of refractory DOM. The presence of oxidized carotenoids in bulk SIO DOM was supported with two-dimensional NMR spectroscopy. Primary support came through the identification of resonances consistent with highly oxidized carotenoids. NMR based estimates for degraded carotenoids in SIO DOM are greater than reduction-based estimates, at approximately, 50%. The demonstrated accumulation of degraded carotenoid compounds in both terrestrial (SRFA) and marine (SIO) DOM is a major contribution of my thesis. It validates other studies which imply compositional similarity between terrestrial and marine DOM, in spite of isotopic differences. One potential for further work would be to make compound specific radiocarbon measurements on the alicyclic reduction products. While attractive, this would require much more work, as the potential for radiocarbon contamination is significant. Other conclusions and future directions will be covered in Chapter IV.

#### *Chapter IV*

In this chapter, I tested whether carotenoid degradation products could be produced over relatively short time-scales under controlled conditions. To that end, I conducted several controlled experiments with the model carotenoid  $\beta$ -carotene. I oxidized  $\beta$ -carotene in both freshwater (MilliQ) and filtered seawater, finding similar

oxidation products in both treatments. In freshwater oxidations I recovered 2.7% of the starting compounds as water soluble oxidation products after 18 days. The compounds present at the end of the experiment were compositionally much more diverse than the starting material, as evidenced by broad peaks in the  $^1\text{H}$  NMR spectrum. The GCxGC-TOF-MS analysis following derivatization of the DOM isolated at the end of the experiment revealed a broad distribution of compounds similar to that observed in derivatized environmental samples. A much greater conversion was observed in seawater oxidations. After 24 days, 6.6% of the carbon was recovered as water-soluble products, and products from both experiments appeared to be identical by NMR and GC-MS. The difference in conversion is in part attributed to the presence of an emulsion layer in the freshwater oxidations that is not present in the seawater oxidation. After sufficient material was collected, the degradation products were chemically reduced. The reduction products were mainly a series of alicyclic hydrocarbons identical to those that had been previously observed in reduced environmental samples (e.g. SRFA, SIO DOM). In addition to previously observed peaks corresponding to the two reduced model compounds observed in Chapter III, prominent evenly spaced peaks were observed ranging from C10-C14, consistent with the model for carotenoid oxidation that was presented in that chapter. These compounds likely correspond to the most prominent series of degradation products formed through oxidation of the isoprenoid backbone. The presence of every compound in this series (C<sub>10</sub>-C<sub>14</sub>), without notable changes in abundance is surprising. While I observed that certain fragments were much more prominent than others, I was not expecting to find the muted odd-even difference in abundance.

The work conducted in this chapter leads into the first major area for future work, which is the mechanistic understanding of carotenoid degradation and subsequent DOM formation in aqueous systems. This work would necessarily include the in-depth structural characterization of the water-soluble  $\beta$ -carotene degradation products. High-resolution mass spectrometry of the degraded compounds would also be useful to compare to environmental samples. Finally,  $\beta$ -carotene is only one out of many marine carotenoids. Further work is necessary to demonstrate how prevalent carotenoid degradation is in the environment.

### *Chapter V*

In this study solid-phase extracted (SPE) DOM isolated across a salinity transect was chemically characterized using bulk and targeted analytical techniques, some of which have never before been applied to DOM analysis. In bulk elemental and isotopic analysis of the sample set, I observed trends largely consistent with conservative mixing between terrestrial and marine DOM pools. However I also observed increases in the N/C elemental ratio and  $^{13}\text{C}/^{12}\text{C}$  ratio that was consistent with mid-salinity organic matter production of nitrogen rich compounds. This deviation from mixing was also observed in DOC and DON concentrations. Derivatization of bulk and hydrolyzed samples was used to semi quantitatively compare samples. FID quantification of different regions within the chromatogram yielded useful information about the distribution of small versus large compounds. I also identified quantifiable compositional differences between Low salinity (terrestrial) and High salinity (marine) samples. Mid salinity samples, expected to be a mixture, were decidedly more similar to Low salinity samples. The dataset

indicated a shift in DOM composition that occurred in the mid-salinity range, which has been observed in other studies as well (Abdulla et. al. 2013). Analysis of the sample set by GCxGC-TOF-MS showed an overall increase in small compounds, and specifically carbohydrate compounds, as salinity increased. Overall, ratios of observed amino acid and sugar monomers did not change; the prominent exception was fucose, which increased in mid-salinity environments. Of the 100s of identified peaks, only the ~20 for which standards were available could be identified. Compounds within “sugar” regions shared mass spectra highly similar to standard compounds, but could not be definitively identified. This is one area for immediate future work. Based on retention times, many of these peaks are hypothesized to be methyl sugars, for which standards are currently unavailable (but have been synthesized in the literature, Panagiotopoulos et. al. 2007). Finally, select samples were chemically reduced. In all samples, reduction products similar to those in Chapters III and IV were identified. Direct comparison of reduction products showed that Low samples are distinct from High samples in distribution, and that each is highly similar to respective salinity endmembers (SIO and SRFA). The one mid-salinity representative had reduction products that were more similar to low salinity reduction products than high salinity reduction products. This is in relative agreement with GC-FID data that indicated that low and mid salinity compounds were compositionally similar. These results suggest DOM composition is not tightly linked to salinity – as would be expected if conservative mixing dominated DOM composition - and that other factors, such as production in mid-salinity waters, could be significantly altering DOM composition. This sample set was also analyzed by collaborators, and

would strongly benefit from those additional data sets. These will be incorporated into the manuscript prior to publication.

## References

- Abulla, H. A. N., Minor, E. C., Dias, R. F., Hatcher, P. G. 2013. Transformations of the chemical compositions of high molecular weight DOM along a salinity transect: Using two dimensional correlation spectroscopy and principal component analysis approaches. *Geochim. Cosmochim. Acta* 118: 231-246
- Hansell, D.A. 2013. Recalcitrant dissolved organic carbon fractions. *Annual review of marine science* 5: 421-445.
- Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., Kettrup, A., Hedges, J.I. 2006. Characterization of a major refractory component of marine dissolved organic matter. *Geochim Cosmochim Acta* 70: 2990-3010
- Lam, M., Baer, A., Alae, M., Lefebvre, B., Moser, A., Williams, A., Simpson, A.J. 2007. Major structural components in freshwater dissolved organic matter. *Environmental Science and Technology* 41: 9240-9247
- Leenheer, J.A., Nanny, M.A., McIntyre, C. 2003. Terpenoids as major precursors of dissolved organic matter in landfill leachates, surface water, and groundwater. *Environmental Science and Technology* 37: 2323-2331.
- Leenheer, J. A., Rostad, C.E. 2004. Tannins and terpenoids as major precursors of Suwannee River Fulvic Acid. U. S. Geological Survey Investigations report 2004-5276.
- Panagiotopoulos, C., Repeta, D. J., Johnson, C.J. 2007. Characterization of methyl sugars, 3-deoxysugars and methyl deoxysugars in marine high molecular weight dissolved organic matter. *Organic Geochemistry* 38: 884-896
- Stenson, A.C., Marshall, A.G., Cooper, W.T. 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: 1275-1284.