CHAPTER 2
Chemical characterization and cycling of dissolved organic matter

NON-PRINT ITEMS

Abstract
This chapter summarizes advances in our knowledge of dissolved organic matter (DOM) composition, with a particular emphasis on studies completed over the last decade that utilize high field nuclear magnetic resonance, high resolution mass spectrometry, proteomics and related immunochemical assays. Up to 75% of marine dissolved organic matter can now be recovered for analysis using solid phase extraction, ultrafiltration, or electrodialysis/reverse osmosis. Spectral and chemical analyses show carbohydrates, proteins, and structurally complex carboxyl-rich aliphatic matter (CRAM) contribute the majority of characterized material. The chemical composition of DOM has a large impact on its accumulation and residence time in the ocean. Carbohydrates and proteins are cycled much more quickly, have much shorter residence times and lower global inventories than CRAM. Processes that lead to the accumulation of carbohydrates and proteins are highly selective, and only a specific fraction of these biopolymers escape degradation to accumulate as DOM. The origin and fate of CRAM are largely unknown, although recent studies suggest a portion of this material has been subjected to elevated temperatures, either in hydrothermal systems or during biomass burning of terrestrial organic matter before transport to the ocean. Radiocarbon measurements and global surveys of deep sea DOM concentration are providing new insights into the location and timescale of refractory DOM removal processes, significantly enhancing our understanding of CRAM cycling.

Key Words
Dissolved organic matter; composition; cycling; marine; polysaccharides; proteins; humic substances; ultrafiltration; solid phase extraction; microbial cycling; photochemistry

Chapter starts here
2.1 Introduction

Each year, between 15-25 Pg of dissolved organic matter (DOM) are added to seawater by the activity of marine microbes, by atmospheric, fluvial, and groundwater transport of organic matter from the continents, and by the release of organic matter from the benthic boundary layer (Burdige, 2007; Jurado et al., 2008; Bauer and Bianchi, 2011; Hansell, 2012). Most DOM is immediately respired by marine microheterotrophs, oxidized by photochemical processes, or permanently buried in sediments. However, a significant fraction is stored in the water column where it interacts with a variety of
biogeochemical cycles over timescales ranging from hours to millennia. Marine DOM
stores nitrogen and phosphorus that would otherwise be immediately available to
microbes in the upper water column, affects the bioavailability of essential trace metals,
attenuates the penetration of UV and visible light in the euphotic zone, and sequesters an
amount of carbon approximately equal to atmospheric carbon dioxide. All production,
removal, and transformation processes leave an imprint on the composition of DOM, and
it is the potential that DOM composition can help to understand the sources and sinks of
DOM, and how it is cycled in the water column, as well as the interest in describing one
of the largest reservoirs of organic matter on Earth, that drives much of the current
research on DOM composition.

Composition refers to the broad suite of molecular characteristics that define
DOM, and includes levels of detail that range from simple elemental ratios of carbon,
nitrogen, and phosphorus, to stable and radio-isotopic content, to the stereochemistry of
amino acids. How DOM is characterized, the extent to which the composition is known,
therefore changes with the characteristics of interest. Global surveys of dissolved organic
carbon (DOC), nitrogen (DON) and phosphorus (DOP) completed over the last decade
now include > 10^4 analyses, and our knowledge of DOM elemental composition is
therefore relatively comprehensive (Hansell et al., 2009, 2012). In contrast, our
understanding of DOM molecular composition is often benchmarked as the inventory of
dissolved compounds (simple sugars, amino acids, lipids, vitamins, pigments, etc.) that
can be isolated from seawater. Viewed from this perspective, < 10% of DOM has been
characterized, and it is fair to say that our knowledge of DOM has improved only
marginally over the last decade and a half. However, it is often other features of
composition, such as the distribution of major carbon, nitrogen and phosphorus
functional groups, the identity of major classes of compounds that contribute to labile and
refractory fractions of DOM, or changes in the degree of oxidation, that are most
informative to understanding the sources, sinks, and cycling of DOM. Viewed from this
perspective, our understanding of DOM composition has advanced significantly to the
point where between 60-70% of DOM has now been “characterized”.

This chapter summarizes recent advances in our knowledge of DOM composition,
with a particular emphasis on studies completed over the last decade that utilize high field
nuclear magnetic resonance, high resolution mass spectrometry, proteomics and related immunochemical assays. A number of excellent reviews and workshop reports summarize the field of DOM composition up to the early 2000’s (McNichol and Aluwihare, 2007; Mopper et al., 2007; Aluwihare and Meador, 2008), and other chapters in this book describe progress in DOM elemental composition (Carlson and Hansell chapter), isotopic composition (Beaupre chapter), dissolved organic nitrogen (Sipler and Bronk chapter) and phosphorus (Karl and Björkman chapter), chromophoric DOM (Mopper et al. chapter, Stedmon and Nelson chapter), and DOM cycling in river and coastal systems (Raymond and Spencer chapter). These topics are integral to DOM composition and cycling, and specific aspects are included in the discussion below, but the reader is referred to these chapters for comprehensive discussions on these topics.

In practical terms, all studies of DOM composition begin with sampling, which aside from providing the material used in chemical and spectral analyses, also selectively defines the chemical fraction and spatial/temporal features of the DOM that is characterized. Because sampling is so integral to the interpretation of DOM composition, this chapter begins with a summary of sampling methods. From there, the chapter is organized into discussions of carbohydrates, proteins and aliphatic organic matter, the major components of DOM that have been identified to date. Finally, the results from these studies are discussed within the broader perspective of how composition impacts the cycling of DOM in the water column.

2.2 Isolation of dissolved organic matter from seawater.

Dissolved organic matter is operationally defined as the fraction of organic matter not retained by filtration. However, the specific choice of filter is determined by the preferences of the analyst and may be influenced by the subsequent suite of analyses that are to be performed. There is no universally agreed upon filter type or pore size that distinguishes dissolved and particulate phases. Many studies of marine particles have utilized Whatman GF/F glass fiber filters with a nominal pore size of 0.7 µm due to the ease with which these filters can be cleaned and to their excellent flow characteristics, facilitating their adoption for DOM studies where large volume samples were often required. However, some marine bacteria and viruses are < 0.7 µm and pass through
GF/F filters to be included in the “dissolved” fraction. Membranes with smaller pore sizes (0.1-0.2 µm) effectively exclude bacteria, while much smaller pore sizes (10-15 nm) are needed to exclude viruses. “Dissolved” is therefore a non-specific term that can include bacteria-sized particles, colloidal organic matter, and truly dissolved species.

Although the concentrations of some organic compounds (amino acids, simple sugars, low molecular weight acids, ketones, and aldehydes) can be measured directly in seawater, > 90% of DOM needs to be concentrated and isolated before spectral and further chemical characterization can proceed. Separation of DOM (~ 0.5-1 mg/L) from salt (~35 g/L) poses the major challenge for DOM isolation, and no single approach recovers all DOM from seawater. Therefore, a number of different strategies have been developed that capitalize on either the larger molecular size or the lower polarity of DOM relative to sea salt. Advances in spectroscopic and spectrometric analyses have in many cases reduced sample requirements from mg to µg or even ng amounts of material, and have mitigated some of the need to obtain DOM as a salt free preparation. For example, samples for mass spectral analyses can be obtained from only a few liters of seawater, and proton nuclear magnetic resonance (1H NMR) spectra have now been reported for unprocessed seawater (Dittmar et al., 2008a; Lam and Simpson, 2008). Presently, these approaches are limited to qualitative analyses of DOM; water suppression leads to biases in the determination of carbohydrate functional groups in 1H NMR, while matrix effects and differences in ionization efficiency complicate the quantitative interpretation of mass spectra. As analytical techniques continue to advance, these limitations will be addressed, but for the immediate future, our ability to isolate only a portion of DOM will continue to limit our understanding of DOM composition.

2.2.1 Isolation of hydrophobic DOM by solid phase extraction

Simple passage of filtered seawater across a solid hydrophobic surface or mineral phase leads to the adsorption and concentration of DOM (Fig. 2.1). Since the basis of the method is physisorptive attraction between DOM and the solid phase, the approach selectively concentrates the hydrophobic or surface-active fraction of DOM with distinct characteristics that are not representative of total DOM. For example, colored dissolved organic matter is efficiently extracted by contact with polystyrene or octadecyl-silica (C-
resins, but the absorption and fluorescent properties of the extracted samples are significantly different from the original seawater (Green and Blough, 1994). Early work on DOM solid phase extraction (SPE) explored charcoal, freshly precipitated iron oxy-hydroxides, and synthetic hydrophobic resins as substrates, but over the past two decades commercially available octadecyl-bonded silica (C-18), cross-linked polystyrene (XAD-2, -4, and -16) and their derivatives (PPL, Isolute ENV, and polyacrylate (XAD-8)) have become the sorbents of choice for these studies (Mopper et al., 2007).

To maximize sample recovery, filtered samples are often acidified with hydrochloric acid to pH 2-2.5. At lower pH, most carboxylic acids and phenols are protonated, reducing their aqueous solubility and enhancing their adsorption, thereby leading to a higher recovery of carbon. The extraction efficiency will depend on the speed at which seawater is passed through the column, and the amount of sorbent used per volume of seawater. A recent study used 1g sorbent per 1-2 mg of total DOC at 20 column volumes per minute, but the effects of flow rate and sorbent/sample ratio have not been investigated in detail (Dittmar et al., 2008b). After collection, the sample is rinsed of salt using pH 2, ultra-pure water, and recovered by elution with methanol and/or methanolic or aqueous sodium or ammonium hydroxide. In performing SPE, the sample is subjected to large changes in pH, which have unknown effects on the sample composition.

The diversity of SPE products available to the analyst has led to a few comparisons of carbon recovery and the elemental, spectroscopic and isotopic characteristics of SPE-DOM (Macrellis et al., 2001; Dittmar et al., 2008a). DOC and analyte recovery was highly dependent on the choice of solid phase. In one study, PPL (functionalized polystyrene/divinyl benzene) was found to be more efficient than C-18 (octadecyl-functionalized silica with 500Å pore size) in extracting DOC from surface seawater near the North Brazilian Shelf, adsorbing on average 62% of total DOC, with a lower C/N (20) and less depleted $\delta^{13}$C value (-23.4‰) than recovered by C-18 (37%, 37 and -24.8‰, respectively). Given the large difference in recovery, C/N ratio and isotope value it would seem that PPL and C-18 extract different but overlapping fractions of DOM. However, the $^1$H NMR spectra of the two samples were nearly identical. In this particular study, differences in extraction efficiency may have been accentuated by the
presence of terrestrial organic matter sourced from the nearby continental margin.

Extraction efficiencies for open ocean seawater have not been widely reported, but are generally lower (10-20%).

2.2.2 Isolation of high molecular weight DOM by ultrafiltration

“Molecular” or “ultra-” filtration exploits the larger hydrodynamic diameter of high molecular weight DOM compared to most dissolved inorganic species to achieve a separation of organic matter from seawater. In this technique, hydrostatic pressure is applied across a semi-permeable membrane perforated with very small pores, typically 1-15 nm in diameter. Salts, water, and organic matter of a hydrodynamic diameter smaller than the filter pore size pass through the membrane (as the permeate) while high molecular weight organic matter is retained and concentrated in the original sample (as the retentate; Fig. 2.1). Accumulation of organic matter on the filter surface rapidly leads to membrane polarization, which is reduced by applying a second (tangential or cross) flow perpendicular to the direction of filtration. The sample is concentrated to a fraction of its original volume, diluted with ultra-high purity, salt-free water, and filtered again. This “diafiltration” process is repeated until the salt content is reduced to a fraction of the organic matter concentration. Once this has been achieved, the remaining water is removed by lyophilization or rotary evaporation, and the sample recovered.

Ultrafiltration concentrates a hydrophilic, high molecular weight fraction of DOM (HMWDOM). The amount and chemical characteristics of HMWDOM recovered is highly dependent on the membrane material (cellulose, polysulfone, etc.) and pore size, the strength of the cross flow, the ratio of original to final sample volume (concentration factor), and a number of other operational parameters. Plots of water volume filtered against DOC concentration show that ultrafiltration with a 1 nm pore-sized membrane typically retains ~30% of DOC, but losses associated with high concentration factors and diafiltration lead to far lower physical recoveries, typically 10-15% for final isolates that are 35-38% by weight carbon (Walker et al. (2011)). Walker et al. (2011) applied a permeation model to ultrafiltration data where seawater was concentrated by factors of >1000 fold. High molecular weight components of DOM behaved ideally, e.g. even at high concentration factors they are efficiently retained by the ultrafiltration membrane.
Hydrophobic, low molecular weight DOM (LMWDOM) that is nominally smaller than the membrane pore size, is also concentrated by ultrafiltration, probably due to membrane polarization and adsorption onto the membranes themselves. The amount of LMWDOM retained is highly variable, being sensitive to the membrane material, concentration factor and perhaps other operational factors. A number of studies report differences in the recovery and chemical properties of HMWDOM. These differences could be due to the variable retention of low molecular weight components. The selectivity inherent in ultrafiltration sampling, and the sensitivity of DOM recoveries to operation parameters that differ between systems and operators, suggests caution must be exercised when comparing data between studies.

2.2.3 Isolation of DOM by reverse osmosis/electrically assisted dialysis

A recent advance in DOM sampling is reverse osmosis/electrically-assisted dialysis (RO/ED; Vetter et al., 2007; Gurtler et al., 2008; Koprivnjak et al., 2009). In this technique, the sample is desalted by an alternating series of positive and negative ion exchange membranes under the influence of an electric potential (Fig 2.1). Anions pass through positively charged ion exchange membranes towards the anode, while cations pass through negatively charged ion exchange membranes towards the cathode. The resulting lower salinity sample is reduced in volume by reverse osmosis, then desalted a second time in a final ED phase. Recoveries of DOM range from 50 to >100% with an average carbon recovery of 76% (n=21), similar to the 70% recovered when SPE and ultrafiltration are used in series (Simjouw et al., 2005). Final preparations are ~25% salt by weight and have chemical characteristics of both hydrophobic organic matter isolated by SPE and hydrophilic DOM isolated by ultrafiltration.

2.3 Chemical characterization of dissolved organic matter

DOM includes simple biochemicals (amino acids, simple sugars, vitamins, fatty acids), complex biopolymers (proteins, polysaccharides, lignins) and very complex degradation products of unknown origin that so far have defied full characterization (humic substances, black carbon). The complexity and diversity of organic constituents
in DOM have pushed the limits of the chemical and spectral techniques brought to bear on their characterization. High resolution mass spectrometry, high field NMR, and new approaches to chemical degradation have significantly broadened and deepened our understanding of DOM composition, but have also highlighted the limits of even advanced spectral techniques.

Ultrafiltration, solid phase extraction, ED/RO, and direct chemical analysis of unfractionated seawater all access overlapping, but compositionally different fractions of DOM. NMR and chemical analysis of hydrolysis products have been most successful in characterizing polysaccharides and proteins in HMWDOM. High-resolution mass spectrometry, 2D NMR and chemical degradation approaches have found their most successful application in the characterization of hydrophobic DOM isolated by solid phase extraction. The following discussion is therefore organized around the characterization of carbohydrates, proteins, and hydrophobic humic substances that are the major components of DOM, highlighting the new approaches and methods that have come to the fore in the last decade.

2.3.1 Polysaccharides in DOM

$^{13}$C NMR analysis of surface water HMWDOM gives a characteristic spectrum (Fig. 2.2A) with major resonances assigned to carboxyl/amide carbon ($\sim$180 ppm; 5%), unsaturated C=C/aromatic carbon (broad peak centered at $\sim$140 ppm; 5%), anomic O-C-O ($\sim$100 ppm; 14%), O-alkyl C ($\sim$75 ppm; 56%), a broad peak centered at $\sim$35-40 ppm (13%) from methylene and substituted methylene C, and two alkyl carbon peaks at 20 ppm (5%) and 26 ppm (5%). $^1$H NMR spectra are similarly characteristic (Fig. 2.2B), with major resonances from anomic (5.2 ppm; HC(-O)$_2$), and O-alkyl (3.5-4.5 ppm; HO-CH) protons, as well as methyl carbons from acetamide (2.0 ppm; HN-C(O)CH$_3$) and deoxysugars (1.3 ppm; C(H$_2$)-CH$_3$). In addition, $^1$H NMR spectra have a distinct signal at 2.7 ppm which is tentatively assigned to N-methyl-acetyl (CH$_3$-NH-C(=O)CH$_3$) on the basis of chemical shift comparisons with literature values, but further corroborative evidence is lacking. The proton and carbon NMR spectra are complimentary, each showing a majority of functional groups can be assigned to carbohydrates. The positions of carbohydrate peaks change little between samples, but
changes in their relative intensities are observed with depth (Benner et al., 1992; Hertkorn et al., 2006), sampling location (Aluwihare et al., 1997), and across salinity gradients (Abdulla et al., 2010a,b), supporting the idea of at least two major components to HMWDOM; a polysaccharide fraction referred to as acylated polysaccharide (APS; Aluwihare et al., 1997) or heteropolysacharide (HPS; Hertkorn et al., 2006; Abdulla et al., 2010a,b), which includes anomeric, O-alkyl, amide, and methyl carbon (from acetate and deoxy sugars), and a carboxylic acid-alkyl carbon-rich fraction referred to as carboxyl-rich aliphatic matter (CRAM) or carboxyl-rich compounds (CRC) reminiscent of aquatic humic substances.

The presence of distinct polysaccharide and CRAM/CRC fractions within HMWDOM can be demonstrated by passing HMWDOM samples through hydrophobic C-18 resin at low pH or through anionic exchange resin at neutral pH to selectively remove most of the CRAM/CRC fraction (Fig. 2.3A; Panagiotopoulos et al., 2007), and through correlation spectroscopy (Fig. 2.3B; Abdulla et al., 2010a,b; Abdulla et al., 2013). In correlation spectroscopy, the variable intensities of major resonances are correlated with a second variable (salinity, depth, location) to derive synchronous and nonsynchronous changes in spectral characteristics. Signal fluctuations that are synchronous indicate a common chemical constituent, while nonsynchronous signals indicate chemically distinct components. Correlation spectroscopy therefore identifies functional group relationships within the different fractions of DOM, but requires that different fractions of HMWDOM vary independently across changes in depth, location, salinity, etc. For example, Abdulla et al. (2013a,b) used two-dimensional (2D) correlation analysis on a suite of samples from the Elizabeth River/Chesapeake Bay estuary to identify synchronous changes in $^{13}$C NMR signal intensity. Strong correlations were observed between intense signals from O-alkyl carbon at 74 ppm, anomeric carbon at 103 ppm, amide carbon at 178 ppm, and two alkyl carbons at 20 and 26 ppm. These functional groups arise from the HMWDOM polysaccharide fraction. The carboxyl-rich component (CRAM/CRC) showed correlation between signals in the methylene and substituted carbon region (29-50 ppm), unsaturated C=C/aromatic region (115-160 ppm), carboxyl carbon at 183 ppm, and carbonyl (aldehyde and ketone) carbon at 190-200 ppm. The negative correlation between carbohydrate resonances at 110 ppm and 74 ppm with
aliphatic signals between 30-46 and unsaturated/aromatic signals centered at 130 ppm indicates that carbon functional groups in these regions of the spectrum are primarily associated with CRAM/CRC. HMWDOM polysaccharides have little to no aliphatic component other than carbon associated with acetate and deoxysugar carbon (Fig. 2.2A).

Full characterization of any polysaccharide is challenging, and typically includes acid hydrolysis to determine simple sugar composition, methylation and reductive cleavage to determine branching, and spectroscopic or spectrometric characterization of partially degraded oligosaccharides to determine sequence. For HMWDOM, these approaches have proved to be only partially successful. Hydrolysis of HMWDOM using a wide variety of acids and hydrolysis conditions yields a characteristic suite of seven major neutral sugars including arabinose and xylose (pentoses), glucose, galactose, and mannose (hexoses), fucose and rhamnose, (6-deoxy-hexoses) (Sakugawa and Handa, 1985; McCarthy et al., 1996; Aluwihare et al., 1997; Panagiotopoulos and Sempéré, 2005), and similar amounts of the amino sugars glucosamine and galactosamine (Kaiser and Benner, 2000; Aluwihare et al., 2002; Benner and Kaiser, 2003). Remarkably, the relative proportions of these sugars is largely conserved across samples collected in different ocean basins, at different times, and at different depths(Sakugawa and Handa, 1985; McCarthy et al., 1996; Aluwihare et al., 1997). However, small variations in the ratio of neutral sugars have been attributed to spatial/temporal changes in HMWDOM composition (Boon et al., 1998; Goldberg et al., 2009, 2010). Quantitatively, the seven neutral and two amino sugars represent only a minor fraction, between 10-20%, of the total HMWDOM carbohydrate. The composition of most HMWDOM polysaccharide, even to the level of simple sugar compliment, is therefore unknown.

 Sugars in HMWDOM have also been characterized by direct temperature resolved mass spectrometry (DT-MS; Boon et al., 1998; Minor et al., 2001). In DT-MS, the sample is rapidly heated under vacuum. Thermal decomposition of the sample leads to the release of simple, volatile degradation products that are characteristic of the type of organic matter undergoing thermolysis. DT-MS of model compounds (proteins, lipids, carbohydrates and nucleic acids) allow for the assignment of diagnostic masses for different compound classes. DT-MS spectra of HMWDOM display ions for hexoses, pentoses, and deoxysugars, that were attributed to N-acetyl-amino-, and monomethyl-
and dimethy-, and methyl-deoxysugars. Ions from N-acetyl-aminosugars were often the most intense features of the mass spectra, consistent with the assignment of the large signal at 2 ppm in the $^1$H NMR spectrum as -C(=O)CH$_3$ (Fig. 2.2B). No ions for acetic acid were observed, and the study concluded that nitrogen, not oxygen, was the site of acetylation. Masses indicative of deoxy- and methyl-sugars were also prominent, and these sugars make substantial contributions to the uncharacterized portion of HMWDOM. HMWDOM polysaccharides are introduced into the mass spectrometer by thermal desorption, which occurs in two distinct stages. The bimodal thermal evolution profile is indicative of at least two fractions of carbohydrate (Boon et al., 1998). Both fractions include the same suite of sugars, but the relative intensity of ions derived from furfural (from uronic acids or deoxysugars) was enhanced and the specificity of the spectra decreased in the high temperature fraction. No evidence was found in either fraction for extended homopolysaccharides: chitans, glycans, xylans, or arabino-galactans that are typical of many storage and some structural polysaccharides. DT-MS data suggest HMWDOM polysaccharides are very heterogenous, include a large fraction of deoxy- and N-acetyl aminosugars, and are highly branched and cross-linked.

Branching into two and three-dimensional polysaccharides can be assessed by selective permethylation of the polysaccharide to protect non-linked sites, followed by hydrolysis. Hydrolysis leaves methylated sites intact, but results in a suite of simple methylated sugars that can be identified and quantified by chromatography. Linkage analysis shows that of the sugars that can be recovered after acid hydrolysis (e.g., 10-20% of the total polysaccharide), 40% have only terminal linkages, 40% are linked without branching, and 20% are branched with one branch point. All of the seven major neutral sugars characteristic of HMWDOM (glucose, galactose, mannose, fucose, rhamnose, xylose and arabinose) have terminal linkages, but only a few sugars are branched. Major linkage patterns for non-branched sugars include 1,3 and 1,4 linkages, with only small amounts of 1,2 branching (Sakugawa and Handa, 1985; Aluwihare et al., 1997). The degree of branching in HMWDOM polysaccharides is unusually high. Highly branched polysaccharides are typical of structural biopolymers found in microbial cell walls. The high degree of branching may impart some resistance to microbial degradation that persists and allows these polymers to accumulate as DOM.
The distinct peaks at 2.0 ppm in $^1$H NMR and 26 ppm in $^{13}$C NMR spectra (Fig. 2.2 A, B) are assigned to the methyl group of N-acetylated sugars (Aluwihare et al., 2005; Quan and Repeta, 2007). Correlation analysis (Fig. 2.3B; Abdulla et al., 2010a,b) shows synchronous changes in resonances in amide carbon at 180 ppm, anomeric carbon, O-alkyl carbon and methyl carbon at 26 and 20 ppm, in support of this assignment.

Based on integration of the $^{13}$C- and $^1$H NMR spectra (Fig. 2.2), acetate contributes 5-8% of the polysaccharide carbon, larger than any other molecular component of HMWDOM polysaccharide identified to date. On this basis, approximately one out of every four sugars in HMWDOM polysaccharide would have N-linked acetate. Aluwihare et al. (2005) inferred that N-acetyl amino sugars are incorporated into a family of related acylated polysaccharides (APS) that includes major neutral sugars and amino sugars, but correlation analysis suggests a somewhat higher ratio (6:1) of neutral to amino sugars, and that there are at least two distinct fractions of neutral and amino polysaccharides.

Correlation analysis distinguishes acylated amino sugar containing polysaccharide (APS) from (non-acylated) heteropolysaccharides on the basis of changes in the asymmetric amide stretching the IR spectrum at 1660 cm$^{-1}$. In a suite of samples through a coastal estuary, Abdulla et al. (2010a) found that as HMWDOM samples become increasing marine, with a relative increase from 50% to 70% in heteropolysaccharide content but little change in the acylated amino sugar containing polysaccharide component, suggesting that the major heteropolysaccharide and aminosugar components of HMWDOM behave independently. A more detailed characterization of HMWDOM and a better understanding of spatial/temporal changes in HMWDOM cycling will be needed to reconcile these two interpretations of polysaccharide composition.

The contribution of N-acetyl-aminosugars to HMWDOM polysaccharide has been quantified indirectly by monitoring the effects of mild acid hydrolysis on $^1$H- and $^{15}$N-NMR spectra (Aluwihare et al., 2005). $^{15}$N NMR spectra of HMWDOM are characterized by a large peak from amide-N at 124 ppm and a smaller peak at 35 ppm from methyl- and amino-N (McCarthy et al., 1997; Aluwihare et al., 2005). The two common classes of biochemicals most likely to contribute to amide-N are proteins, polymers of amino acids linked through an amide functional group, and N-acetyl amino sugars such as chitin or peptidoglycan. Hydrolysis of either will convert amide-N to
amino-N, however for proteins this results in depolymerization while in N-acetyl amino
sugars it does not. If acetate occurs primarily as N-acetyl amino sugars, then the amount
of amino acids plus acetic acid released by hydrolysis should equal the conversion of
amide-N to amino-N measured by $^{15}$N NMR. If conversion exceeds the sum of amino
acids and acetic acid, then other biochemicals must contribute to amide-N. Alternatively,
if the conversion is significantly less than the sum of amino acids and acetic acid, then a
significant fraction of acetate is bound as (O-linked) esters.

Hydrolysis of surface (5-23 m) and midwater (600-1000 m) samples showed good
agreement between the conversion of amide-N to amino-N and the production of amino
acids and acetic acid (Fig. 2.4). Between 97-116% of the new amino-N could be
accounted for as the sum of amino acids and acetic acid. Of this, the majority (72-90%)
was attributed to hydrolysis of N-acetyl amino sugars, while the balance (11-44%) was
attributed to protein hydrolysis. Critical to this interpretation is the assumption that
quantitative agreement between $^{15}$N NMR and molecular level acetic acid analysis
signifies that acetic acid is derived from N-acetyl amino sugars. The experiment does not
show this directly, but given results from DT-MS that only acetamide is generated by
HMWDOM thermolysis, fortuitous agreement between the two techniques seems
unlikely.

Detailed characterization of methyl sugars followed their detection in DT-MS
spectra by nearly a decade. Hydrolysis of HMWDOM yields an O-methylated sugar
fraction that can be purified by chromatography and characterized in detail by $^1$H NMR
and mass spectrometry (Panagiotopoulos et al., 2007; Panagiotopoulos et al., 2013). All
seven major neutral sugars recovered by hydrolysis were found to have mono- and di-O-
methylated homologues (Fig 2.5). In addition, many novel O-methyl sugars not yet
identified in the nonmethylated fraction, including O-methyl heptose, O-methyl di-deoxy
hexoses and yersiniose, a dideoxy-4-C-(1-hydroxyethyl)-D-xylo-hexose, occur in small
amounts. The presence of methyl sugars does not impact the interpretation of linkage
analysis, since methylation prevents linkage through a particular carbon, and the product
in each case is a methyl sugar. O-methyl sugars are common in algal and bacterial
structural polysaccharides, and as such have little biomarker value, but their occurrence
in HMWDOM and the complexity of the O-methyl sugar mixture suggests a broad suite
of structural polysaccharides may co-exist in HMWDOM. Treatment of HMWDOM with periodate at high temperature yields methanol as a major oxidation product (Quan and Repeta, 2007). Assuming all this methanol is sourced from O-methyl sugars, O-methyl sugars contribute an important fraction of HMWDOM, beyond the amount released by acid hydrolysis. Quantitative analysis of methanol after periodate oxidation could be useful in assessing the contribution of methyl sugars to non-hydrolyzable HMWDOM polysaccharide.

Attempts to further characterize HMWDOM polysaccharides have been met with limited success. For reasons that are still unknown, even aggressive hydrolysis by strong acid does not appear to depolymerize HMWDOM, and ~70% of the polysaccharide fraction remains relatively uncharacterized (Panagiotopoulos and Sempéré, 2005). Information on the non-hydrolyzable portion of HMWDOM polysaccharide can be obtained from 2D homo- and heteronuclear NMR correlation techniques (Fig. 2.6). For example, the $^1$H NMR correlation spectroscopy (COSY) spectrum of non-hydrolyzable HMWDOM polysaccharide shows strong cross peaks between methyl, H-6 protons (CH$_3$-) and carbohydrate H-5 (HC-OH) from 6-deoxysugars. Cross peaks are also observed between 2-3 ppm and 3-4 ppm from 2-, 3- or 4-deoxysugars (Fig. 2.6). Deoxysugars are recovered from HMWDOM hydrolysis products only in low yields (~2% of total carbon), but the 2D NMR suggests a much higher contribution to HMWDOM carbohydrate.

Each class of sugar (hexose, pentose, deoxyhexose, aminohexose, methylhexose) identified by DT-MS has a different molecular weight easily distinguished by mass spectrometry, making this an attractive approach for understanding the sequence and arrangement of sugars within DOM (Boon et al., 1998; Minor et al., 2001). However, a number of issues including the low ionization efficiency of carbohydrates relative to other classes of organic matter, incomplete removal of inorganic salts, and a molecular weight range that may easily exceed the mass range of most spectrometers (typically 2-4 kDa), need to be overcome before MS can be applied to its full potential (Sakugawa and Handa, 1985; Schmidt et al., 2003). For example, high-resolution mass spectra for HMWDOM show no ions with H/C and O/C elemental ratios characteristic of polysaccharides (~1.7-2.0 and ~ 0.8-1.0 respectively; Hertkorn et al., 2006). Numerous
strategies have been developed in glycochemistry to enhance the MS detection of polysaccharides by conjugation with proteins or fluorescent tags that are amenable to ionization by electrospray and matrix assisted laser desorption ionization (ESI; MALDI). These approaches might prove promising for DOM polysaccharides, and could facilitate MS characterization. Separation of HMWDOM polysaccharides into defined fractions could also facilitate spectral characterization. Appropriate methods for size exclusion and ion chromatography for complex mixtures of oligo- and polysaccharides in the molecular weight range of a few kDa and higher are limited, but have been applied to HMWDOM (Sakugawa and Handa, 1985). Recent advances in strong anion exchange chromatography at high pH also shows promise as an approach for HMWDOM carbohydrate separations (Corradini et al., 2012).

Finally, some inferences as to the composition of different HMWDOM polysaccharides can be drawn from changes in the temporal and spatial distribution of sugars between samples. Multivariate analysis of DT-MS data from samples collected in the U.S. Mid Atlantic Bight between 35°N-43°N and the Gulf of Mexico find the same suite of ions from hexoses, pentoses, deoxy-, N-acetyl-amino- and methyl sugars in all samples, but the relative abundances changed with sample location (Boon et al., 1998). These differences were attributed to changes in the composition or relative abundance of different polysaccharides. The most striking differences were between sugars desorbed at low and high temperatures. The high temperature fraction from all samples showed a remarkable similarity in composition, while the sugar distribution in the low temperature samples showed much more diversity, and some clustering by sampling site. The results suggest two distinct fractions of polysaccharides co-exist in coastal seawater, a fraction with highly similar composition that is ubiquitous at all sites, and a fraction of variable composition influenced by local production.

2.3.2 Proteins and amino acids in DOM

Proteins account for up to 50% of the organic carbon and 80% of the organic nitrogen in marine microbes. Grazing, release of extracellular enzymes, and viral lysis all introduce proteins and amino acids into seawater, and on this basis alone, it is likely that proteins contribute to DOM. Analytical methods to identify and quantify dissolved
proteins directly (proteomics) have only recently reached the point where they can be 
applied to DOM, and the next decade should see a rapid expansion in our understanding 
of the sources, cycling, and fate of proteins as these methods become more widely 
applied. However, studies of the distribution, stereochemistry, and isotopic values of 
amino acids released after DOM hydrolysis have already made major contributions to our 
understanding of dissolved protein cycling, and have set the stage for the future 
application of proteomics.

Sensitive methods for the detection of nanomolar concentrations of dissolved 
amino acids were first applied to seawater in the early 1960’s and 70’s (Lee and Bada, 
1975, 1977). Due to the low ambient concentrations of dissolved amino acids, and the 
potential for contamination by laboratory glassware and reagents, early studies gave 
variable results. Current analyses capitalize on the reaction of ortho-phthalaldehyde 
(OPA) with primary amines in basic, aqueous solutions to form fluorescent, hydrophobic 
products that can be retained and separated by high pressure liquid chromatography 
(Lindroth and Mopper, 1979; Mopper and Lindroth, 1982). The reaction proceeds 
rapidly and at high yields and has sub-nanomolar detection limits. Primary amines other 
than amino acids (ammonia, urea, etc.) also react with OPA and are included in the 
analysis. Some amino acids co-elute under some separation conditions, but with 
appropriate calibration the method provides good quantitative measurements of dissolved 
amino acids in seawater (Tada et al., 1998).

Dissolved amino acids are operationally classified by the methods used for 
sample processing. Dissolved “free” amino acids (DFAA) are measured by the direct 
reaction of OPA with seawater, and are thought to represent monomeric amino acids 
present in a sample. In order to minimize contamination during sample processing, some 
early studies did not use filtration to separate dissolved and particulate free amino acids 
(Mopper and Lindroth, 1982). However, subsequent studies typically include a filtration 
step (Fuhrman, 1987). Treatment of filtered seawater with acid hydrolyzes peptides, 
proteins, and glycoproteins, and allows for the measurement of “total hydrolysable” or 
“total dissolved” amino acids (THAA; TDAA, respectively; Lee and Bada, 1975, 1977).
The difference; THAA-DFAA represents dissolved combined amino acids (DCAA), the 
fraction of amino acids bound as proteins, peptides and other amino acid polymers. Since
DFAA are typically more than an order of magnitude less abundant than DCAA, recent studies have focused only on total hydrolyzable or total dissolved amino acids (McCarthy et al., 1996; Yamashita and Tanoue, 2003b).

Open ocean profiles of DFAA in the equatorial Pacific and Sargasso Sea show very low values throughout the water column. Surface water concentrations (10-40 nM) are somewhat enriched relative to deep sea samples, which have lower and more constant values (<10-20 nM). THAA distributions follow a pattern similar to DFAA, with higher and more variable concentrations in surface waters (200-450 nM) and lower and more stable values below the euphotic zone (100-200 nM)(Fig. 2.7; Lee and Bada, 1975, 1977; McCarthy et al., 1996; Yamashita and Tanoue, 2003b; Kaiser and Benner, 2009) Semi-enclosed seas, coastal and near-shore sites have higher concentrations of both DFAA and THAA.

Overall, spatial patterns of THAA concentrations reflect changes in DOC concentrations. THAA carbon to total DOC ratios (THAA-C/DOC) vary between sites and depth, but generally fall within 1-4% for surface waters and 0.4-0.8% at depths > 1000 m (McCarthy et al., 1996; Yamashita and Tanoue, 2003a; Kaiser and Benner, 2009). THAA contribute a larger portion of DON (1.4-11%; McCarthy et al., 1996; Tada et al., 1998; Kaiser and Benner, 2009), and are the largest component of DON characterized to the molecular level. A decrease in the ratio of THAA-C to DOC has been observed between highly productive coastal waters and the open ocean along a line extending from Japan (Yamashita and Tanoue, 2003a). THAA carbon to DOC ratios fall from 4% in near-shore surface waters to 2% in offshore surface waters. Similar patterns have been observed in the Baltic Sea, Chesapeake Bay, Biscaye Bay (Florida) and in the Laptev Sea (Mopper and Lindroth, 1982). Nutrient and chlorophyll concentrations were quite high at all sites and higher rates of local primary production; more dynamic organic matter cycling closer to shore may result in relatively higher contributions of THAA to DOC.

Major amino acids include glycine, alanine, glutamic acid, serine, aspartic acid, arginine and threonine, which typically contribute >90% of THAA (Fig. 2.8). Other protein amino acids individually represent < 5 mole % of THAA, although contributions from valine and leucine are sometimes in excess of this amount. Tryptophan decomposes
under the acid and temperature conditions typically used for THAA analysis, and can be measured in samples only when alkaline conditions are used for hydrolysis. In open ocean waters tryptophan contributes ~1 mol % of THAA, although somewhat higher contributions were noted near-shore (Yamashita and Tanoue, 2003b). The distribution of amino acids in THAA is significantly different than particulate organic matter (phytoplankton, suspended and sinking detrital material, Fig. 2.8), with higher concentrations of aspartic acid and glycine, but lower concentrations of arginine, leucine and isoleucine (Fig. 2.8).

Amino acid distributions at open-ocean and coastal sites generally do not show large, systematic changes, however subtle changes have been detected through correlative analysis. Using relative abundance as a measure of sample relatedness, Yamashita and Tanoue (2003b) compared changes in the distribution of amino acids along a north-south transect along 137°E from Ise Bay, Japan into the northwest Pacific Ocean. Some amino acids (for example glycine and alanine) are positively correlated in all samples, and on this basis, amino acids were grouped into four categories. The positive correlations within a particular group indicates similarities in the biogeochemical cycling of the constituent amino acids, while differences in the correlations between groups suggest divergence in THAA cycling due to differences in either the macromolecular form of the peptides or the way THAA are processed by microbial degradation. Amino acid distributions have also been analyzed by principal component analysis (PCA) to quantitatively differentiate patterns of amino acid distribution. Cross correlation of the first and second PCA scores for each acid showed general groupings based on sample location and depth (bay, coast, open-ocean, surface, and deep water), and cross plots of factor coefficients grouped acids into four classes with the same make up as found for correlative analysis.

Amino acids occur in two enantiomeric forms (D- and L-amino acids). Proteins are made exclusively from L-amino acids, but many bacteria utilize D-amino acids in cellular regulatory functions and in bacterial cell wall biopolymers (Cava et al., 2011). Enantiomeric amino acids can be chromatographically separated and quantified in THAA analysis, and the composition and distribution of D-amino acids have been used as a biomarker for the contribution of bacterial cell wall material to DOM (Lee and Bada,
Lee and Bada (1977) reported enantiomeric (D:L) ratios of aspartic acid (0.07-0.44), glutamic acid (0.04-0.07), alanine (0.05-0.14), valine (0.02-0.04), isoleucine (<0.01-0.22) and leucine (0.02-0.07). Subsequent studies have reported D:L ratios for serine of 0.1-1. McCarthy et al. (1998) found the HMWDOM fraction has higher D:L ratios for glutamic acid (0.08-0.15) and alanine (0.32-0.55) than THAA. These ratios are greater than D/L ratios in cellular organic matter and POM, and are indicative of selective preservation of D-amino acid containing bacterial structural polymers relative to proteins in THAA (McCarthy et al., 1998; Dittmar et al., 2001; Nagata, 2003; Kawasaki and Benner, 2006, 2008).

The high relative abundance of D-amino acids in THAA and HMWDOM has led to a series of investigations designed to assess the contribution of bacterial cell wall material to DOM. Bacteria are abundant at all depths in the ocean, and the selective preservation of D-amino acids indicated by high D/L ratios suggests that cell wall biopolymers are more resistant to microbial degradation than proteins, and are therefore likely to be persist in DOM (McCarthy et al., 1996; Kaiser and Benner, 2008; Calleja et al., 2013). Kaiser and Benner (2008) measured the abundance and distribution of D-amino acids and muramic acid (a constituent of peptidoglycan in bacterial cell walls) in marine heterotrophic and cyanobacteria in order to make a quantitative assessment of the peptidoglycan and other D-amino acid containing bacterial cell wall constituents to DOM. In purified peptidoglycan, the ratios of muramic acid to D-glutamic acid and D-alanine are ~1 and ~0.75, respectively. Ratios in marine POM and DOM above these values were interpreted to signify the presence of other (non-peptidoglycan) D-amino acid containing constituents, while ratios below these values signify the preferential degradation of D-amino acids relative to muramic acid. Muramic acid was non-detectable (<1.2 nM) in filtered seawater, even though total D-amino acids concentrations ranged between 8-26 nM. These high relative concentrations suggest that D-amino acids in DOM are either derived from peptidoglycan degradation products or are constituents of bacterial polymers other than peptidoglycan.

Tanoue and colleagues pioneered the application of gel electrophoresis and N-terminal sequencing to characterize intact proteins in DOM (Tanoue et al., 1995; Tanoue, 1995; Tanoue et al., 1996; Yamada and Tanoue, 2003). Dissolved proteins can be
concentrated and desalted using ultrafiltration (>10 kDa), then partially purified by precipitation with cold trichloroacetic acid or methanol/chloroform/water (Tanoue, 1995; Powell and Timperman, 2005). Care must be exercised throughout the concentration steps to minimize contamination from lysed cells or bacteria growth, as well as from losses of proteins due to precipitation or adsorption onto filters and other surfaces. The insoluble protein pellet from trichloroacetic acid precipitation was washed of residual polysaccharides and lipids with ethanol and diethyl ether, then analyzed by gel electrophoresis.

Electrophoretograms of dissolved proteins were compared to a suite of standard marker proteins separated on the gel electrophoresis plate by molecular size (Fig. 2.9). Samples from the northcentral and southwest Pacific Ocean and the Gulf of Mexico show proteins with a wide range of sizes corresponding to molecular weights from < 14kDa to > 66 kDa (Tanoue et al., 1995; Tanoue, 1995; Tanoue et al., 1996; Powell et al., 2005). Major proteins were concentrated in 25-30 bands, some of which were present in all samples, while others changed with sample location and depth. Surface waters <50 m generally showed low numbers and concentrations of recognizable protein bands. The number and intensity of bands increased with depth (75-200m); below 200m, gels were heavily stained, with a high degree of background staining from unresolved proteins and perhaps humic substances. Proteins with putative molecular weights of 48 kDa and 37-40 kDa were clearly visible at many of the stations irrespective of depth, while proteins designated as 66-63 kDa, 44 kDa, 41 kDa, 31-34 kDa, 26 kDa, 23 kDa, and 15 kDa varied with depth and sampling location, but were present in a number of samples. The strikingly similar patterns that appear in electrophoretograms of dissolved proteins led Tanoue et al. (1996) to conclude that the processes that transfer proteins from cellular material and allow for their accumulation in DOM are similar across broad expanses of the ocean.

The major protein band at ~48 kDa present in all samples was recovered from six samples collected between 45-462m and the N-terminal sequence of 14-15 amino acids determined (Tanoue et al., 1995). Later analysis of these samples expanded the sequence to 24 amino acids (Yamada and Tanoue, 2003). All 48 kDa bands yielded the same N-terminal sequence of amino acids, indicating that all bands represented the same protein.
The sequence shared 100% homology with porin-P and porin-O of the gram-negative bacterium *Pseudomonas aeruginosa* (Tanoue et al., 1995; Yamada and Tanoue, 2003). Porin-P and porin-O are membrane proteins expressed to facilitate cross membrane transport of small hydrophilic substrates, often under conditions of phosphate stress. The bacterial community composition at the sampling sites was not determined, but genomic analysis of bacterioplankton at Station ALOHA, one of the stations sampled by Tanoue, yielded 16s sequences of *Alteromonas*, *Vibrio* and *Pseudomonas* spp., bacteria that are closely related to *P. aeruginosa*. Similar biosynthetic pathways between indigenous bacteria and *P. aeruginosa* might lead to porins of comparable homology. Recognizing that existing databases contain sequences from only a small number of proteins, Tanoue’s data suggests that bacterial proteins such as porin-P contribute to DOM. Porin P has been shown to be resistant to proteases, which together with a potentially ubiquitous and abundant source, may explain its appearance in all samples analyzed by gel electrophoresis in Tanoue’s studies. Likewise a 40 kDa protein from the North Pacific had 100% homology with the family of outer membrane proteins (OmpAs) of *Acinetobacter* spp., however sequencing of the 30 kDa, 37 kDa, and 39 kDa proteins did not exhibit homology with any known proteins in searchable databases, and these proteins could not be identified.

Suzuki et al. (1997) and Yamada et al. (2000) developed and applied an immunochemical assay against bovine serum albumin modified with the N-terminal 14 oligomer of *P. aeruginosa* porin P (α-48 DP N-14), and the whole outer membrane protein Omp35La from *Vibrio anguillarum*. Polyclonal antibodies developed against these antigens were used as sensitive screens for porin P and related proteins in DOM. Western blots of DOM proteins showed cross reaction between the α-48 DP N-14 probe and the 39 kDa, 48 kDa and 60 kDa bands, and between the Omp35La probe and the 18 kDa, 34 kDa and and 70 kDa bands. Cross reaction of the α-48 DP N-14 probe and the 39 kDa protein is not fully understood, as a subsequent investigation showed this protein to be glycosylated, and therefore probably not a porin P homologue (Yamada and Tanoue, 2003). Cross reactivity for the α-48 DP N-14 probe was also observed for natural populations of bacteria in all samples, although the number of bacteria cells that cross reacted were 2-6 orders of magnitude less than enumerated by total bacterial cell
counts. The results raise the possibility that the sources and composition of proteins dissolved in seawater might be highly specific.

Erdman degradation was used to partially sequence the 48 kDa protein in DOM, but HPLC-MS techniques offer the possibility of more comprehensive sequencing of dissolved proteins in environmental samples (Powell et al., 2005). In this approach, proteins purified by gel electrophoresis or capillary zone electrophoresis are partially digested by exposure to trypsin, a protease, and the peptides separated by HPLC and sequenced by MS. Fragment ion masses are used to reconstruct the peptide amino acid sequence that is then compared to sequence information stored in databases of known proteins. Using this approach, Powell et al. (2005) distinguished families of proteins, showing that proteins from both bacterial membranes (fatty acid synthetase, luminal binding protein) and enzymes (ribulose bisphosphate carboxylase, anthranilate synthethase) were present in DOM (Powell et al., 2005).

Bottom-up proteomics and amino acid D/L ratios both highlight the contribution of bacterial proteins to DOM. Another way to assess the sources, cycling, and distribution of dissolved proteins is to target abundant proteins using 2D gel electrophoresis and sensitive immunoassays such as ELISA (enzyme-linked immunosorbent Assay) and MSIA (mass spectrometry immunoassay; Orellana et al., 2003; Jones et al., 2004; Orellana and Hansell, 2012). In their proteomic data, Powell et al. (2005) reported short peptide sequences suggestive of ribulose-1,5-bisphosphate carboxylate/oxygenase (RuBisCo). RuBisCo is one of the most abundant proteins on Earth, and is essential in catalyzing carbon fixation in vascular plants, algae, photoautotrophic and chemoautotrophic bacteria. RuBiscCo is abundant in particulate matter located in the euphotic zone, where grazing, viral lysis, and other processes transfer some particulate RuBisCO into DOM. Some of the DOM proteins visualized by electrophoresis and staining, particularly in the range of known RuBisCo subunit proteins of 55 and 13 kD, probably result from this cycling. Orellana and Hansell (2012) used a synthetic protein incorporating RuBisCo sequences to develop an immunoassay for anti-RbcL to measure RuBisCo concentrations in ~800 samples from the North Pacific Ocean. The large number of samples analyzed in this study, and sensitive detection limits (< 1 ng RuBisCo L⁻¹), highlights the potential power of immunological approaches to track
the distribution and cycling of biologically important proteins. They found RuBisCo in all samples from the surface to depths > 4000 m. RuBisCo is synthesized in the euphotic zone, and the presence of high concentrations of RuBisCO (5-20 ng/L) throughout the deep ocean ties meso- and bathypelagic RuBisCo distributions to export production (a small amount of RuBisCO may be synthesized at depth from chemoautrophy, but this contribution is thought to be small). As large particles sink from the euphotic zone, microbial remineralization and physical disaggregation release smaller particles and inject DOM into underlying waters. High relative concentrations (15-20 ng L\(^{-1}\) between 2000-4000 m) of RuBisCO in the deep equatorial and subarctic Pacific coincide with higher carbon export in these regions. Low relative concentrations of RuBisCo (< 10 ng L\(^{-1}\) between 2000-4000 m) below the northern and southern subtropical gyres were attributed to the low carbon export fluxes that characterize these regions. The coupling of deep water RuBisCo concentrations to surface processes implies a rapid turnover of RuBisCo related proteins, but spatial differences in deep RuBisCo concentrations also trace deep water mass flows. If correct, then the residence time of some deep RuBisCo is on the order of years to decades, a timescale typically associated with semi-labile DOM (Section 4.2).

2.3.3 Humic substances in solid phase extractable DOM (SPE-DOM)

Simple \(^1\)H and \(^13\)C NMR spectra along with elemental analysis of hydrophobic SPE-DOM made in the 1970’s and 1980’s showed SPE-DOM is distributed throughout the water column, is rich in COOH and aliphatic carbon, and has a COOH/aliphatic carbon ratio of 1:4-5 (Hedges et al., 1988). Natural products with such a high ratio of COOH/alkyl carbon are rare in nature, and the SPE-DOM fraction is therefore thought to result from extensive transformations of marine lipids, carbohydrates and proteins. Recent studies refer to this fraction of DOM as “carboxyl rich aliphatic material” (CRAM), but it has alternatively been referred to as hydrophobic DOM, solid phase extractable (SPE) DOM, carboxylate rich carbon (CRC), or marine humic substances (HS). For simplicity, this fraction of DOM will hereafter be referred to as SPE-DOM, since all studies rely on isolation by adsorption onto a solid hydrophobic resin (XAD, C-18, PPL), and recent work has begun to distinguish distinct components within SPE-
DOM and assign characteristic molecular features to specific terminology (CRAM, thermogenic DOM, etc.). However, in the literature, different terminologies are still in use. SPE-DOM as been compared to humic substances isolated from soils and freshwaters, but the marine version has a lower aromatic and olefinic content, a higher C/N ratio, and a carbon staple isotope ($\delta^{13}$C) value of $\sim$21-22 $\%$, all of which suggested a autochthonous source (Gagosian and Steurmer, 1977; Hedges et al., 1988; Druffel et al., 1992). Although only a handful of measurements exist, SPE-DOM is highly depleted in radiocarbon (-310 to -587 $\%$), and is considered to represent recalcitrant fractions of DOM resistant to microbial oxidation (Druffel et al., 1992).

2.3.3.1 Characterization of SPE-DOM by high-field nuclear magnetic resonance

The processes that lead to the formation and removal of SPE-DOM in the ocean are not known, but high field NMR and high resolution MS have added unprecedented detail to our knowledge of its composition (Hertkorn et al., 2006; Dittmar and Koch, 2006; Hertkorn et al., 2012; Helms et al., 2013). $^1$H NMR spectra are characterized by broad peaks between 0.8-10 ppm with substantial signal overlap (Fig 2.10). Major signals have been assigned to aliphatic $\text{CH}_3$ and $\text{CH}_2$, carbonyl-rich aliphatics, particularly methine protons ((C)$_2$-$\text{CH}$-$\text{COOH}$), methoxy protons ($\text{CH}_3$-$\text{O}$), carbohydrate-derived methines ((C)$_2$-$\text{CH}$-$\text{OH}$) and (O-$\text{CH}$-$\text{O}$), and protons on sp$^2$ hybridized olefinic and aromatic carbon. When $^1$H NMR spectra of samples collected in a depth profile from 5m to 5446m were normalized to 100% total area (0-10.5 ppm), the spectra showed nearly coinciding aliphatic terminal (-$\text{CH}_2$-$\text{CH}_3$; $\delta$ $\sim$ 0.9 ppm) methyl abundance, variable methylene (-$\text{CH}_2$-$>$ 4 bonds from a heteroatom) abundance, and progressively increasing amounts of H associated with methylated, alicyclic rings with depth. Signals associated with acylated polysaccharides (N-$\text{C(O)}$-$\text{CH}_3$; H-$\text{C-OH}$, O-$\text{CH-O}$, $\text{CH}_3$O-C) decreased with depth.

Although signal overlap limits molecular definition in 1D $^1$H NMR spectra, 2D COSY spectra provide unprecedented definition of overlapping signals. About 4500 off-diagonal cross peaks have been observed for SPE-DOM collected in surface waters, of which $\sim$75% were derived from sp$^3$-hybridized carbon ($\text{H}$-$\text{C}_{\text{sp}}$-$\text{C}_{\text{sp}}$-$\text{H}$, $\text{H}$-$\text{C}_{\text{sp}}$ (O)-$\text{C}_{\text{sp}}$-$\text{H}$, and $\text{H}$-$\text{C}_{\text{sp}}$ (O)-$\text{C}_{\text{sp}}$ (O)$\text{H}$), and 25% from sp$^2$ hybridized carbon ($\text{H}$-$\text{C}_{\text{sp}}$-$\text{C}_{\text{sp}}$-$\text{H}$; $\delta_1$ $>$ 5
ppm; Hertkorn et al., 2012). Assuming an average spin system of ~ 3.5 protons, the COSY spectra of SPE-DOM suggests a mixture of at least several hundred distinct molecular species. With depth, the COSY spectra of SPE-DOM become progressively attenuated, however the position of major peaks remains the same. SPE-DOM composition is highly conserved with depth, but the attenuation of COSY cross peaks suggests subtle changes in composition that lead to faster transverse relaxation of NMR signals and/or increasing molecular diversity associated with DOM aging (Hertkorn et al., 2012).

$^{13}$C NMR spectra and 2D heteronuclear $^1$H/$^{13}$C NMR spectra are consistent with these assignments, allowing for a better quantitative assessment of how carbon functional groups are distributed within SPE-DOM. $^{13}$C NMR spectra also provide some unique insights not available through $^1$H NMR alone. All $^{13}$C NMR spectra show signals from carbonyl (C=O; 220-187 ppm), carboxyl (COX; where X= -O, -N, -CH$_3$; 187-167 ppm), aromatic C-X (where X = O, N) and aromatic C-H (167-145 and 145-108 ppm respectively), carbohydrate carbon (anomeric O-C=H-O, 108-90 ppm and O-alkyl HC-OH, 90-47 ppm), and aliphatic carbon (47-0 ppm; Hertkorn et al., 2006, 2012) The relative amount of carboxylic acids and ketones increases from surface to deep water, and the amount of labile carbohydrate carbon declines, indicating a selective removal of carbohydrates and gradual oxidation of the non-carbohydrate fraction. Methine carbon associated with aliphatic branched functional groups (C$_3$-C-H) increased relative to aromatic and carbohydrate associated carbon from 46% in surface waters to 57% at 5446 m. The highly branched nature of marine SPE-DOM is also indicated by the relatively intense adsorption bands between 2970-2980 cm$^{-1}$ (aliphatic methyl stretch) and 2944 cm$^{-1}$ (aliphatic methyl stretch) in the infrared (Esteves et al., 2009). NMR and IR spectra indicate that aromatic carbon was not abundant at any depth (< 5%), and some aromatic carbon associated with downfield NMR signals ($\delta_c = 164$ ppm) display a cross beak with aromatic protons ($\delta_H = 8.2$ ppm) that indicate nitrogen heterocycles (Esteves et al., 2009; Hertkorn et al., 2012). Only a minor fraction (15%) of methyl groups were purely aliphatic (-CH$_2$-CH$_3$), while a major fraction (70%) were shifted downfield to between 1-1.6 ppm. The downfield shift in most methyl groups suggests proximity (< 3 C bonds away) to carboxyl carbon. The carboxyl carbon peak itself is Gaussian shaped and
displays a sizable chemical shift range (6 ppm at half height) indicating high diversity and little preference or regularity in its chemical environment.

In summary, $^1$H and $^{13}$C NMR spectra show SPE-DOM is a highly complex, yet well defined mixture of molecular components, that includes carbohydrates and a carboxy-rich aliphatic (CRAM) fraction, along with a minor amount of extended aromatic and aromatic N-heterocycles. The CRAM fraction appears to become more highly branched with depth, either from transformations associated with aging, or from selective removal of less highly branched carbon. The major change in the distribution of carbon functional groups in SPE-DOM with depth results from the progressive loss of carbohydrate and the conservation of CRAM (Hertkorn et al., 2012).

2.3.3.3 Characterization of SPE-DOM by high resolution mass spectrometry

High resolution mass spectra of SPE-DOM have been reported since the 1970s (Gagosian and Steurmer, 1977), but the introduction of Fourier Transform Ion Cyclotron Resonance (FT-ICR) and more recently Orbitrap mass spectrometers, each capable of measuring ions at both very high mass accuracy and resolution, has changed our ability to characterize SPE-DOM at the molecular level (Kido Soule et al., 2010; Zubarev and Makarov, 2013). Coupled to electrospray ionization (ESI) these instruments provide the necessary resolution to distinguish the several thousand ions with unique masses within DOM between 200-2000 Da (Fig. 2.11). Due to the high mass resolution and accuracy (typically < 1 ppm for ions < 400 Da), elemental formulas for most ions can be assigned with a high level of confidence. Since the number of isomers for any given mass increases rapidly with molecular weight, each unique mass probably represents a mixture of different isomers. High resolution MS therefore does not allow for the full identification of new compounds, however it has led to the discovery of important new compound classes, and it is proving to be a very powerful technique for detecting changes in DOM composition between samples. To fully exploit the large amount of data provided by high resolution MS, analysts use a number of computational tools to extract information about DOM composition. Typically, elemental formulas with all possible combinations of atoms are calculated and matched to all ions to within a mass precision determined by the operational resolution of the particular instrument.
Depending on the approach used 45-97% of observed masses can often be assigned molecular formulae. Elemental formulae with only C, H, and O dominate the molecular formulae of assigned masses, with fewer elemental formulae assigned to compounds having nitrogen (CHNO), sulfur (CHOS) and both nitrogen and sulfur (CHNOS).

In performing these calculations, some assumptions are needed. First, the type and likely number of atoms within a formula are inferred. For example, in a recent study of SPE-DOM in the southwest Atlantic Ocean, Flerus et al. (2012) constrained molecular formula calculations for masses between 200-600 Da within $^{12}$C (0-∞), $^{13}$C (0-1), $^1$H (0-∞), $^{16}$O (0-∞), $^{14}$N (0-∞), $^{32}$S (0-1). For positive ions $^{23}$Na (0-1) is also allowed to account for sodiated DOM ions. The resulting mass list is filtered to remove $^{13}$C isotopes (mass difference between peaks of 1.003 Da), and further filtered by making assumptions about nitrogen, the H/C ratio, and the number of double bonds (Koch et al., 2007). The power of high resolution mass spectrometry lies in its ability to distinguish between elemental formula with the same nominal masses. For example, the functional group $\text{CH}_3\text{-CH-R}$- has the same nominal mass (16 Da) as the functional group $\text{O=CH-R}$-, but differs in exact mass by 36.4 mDa (16.0313 vs 15.9949), which is well within the resolution of high field instruments. High resolution MS data allow mass lists to be grouped into “pseudo-homologous” series that differ by a specified functionality (e.g., -$\text{CH}_2$- or -$\text{CO}_2$- ). The measured mass is converted to a “Kendrick mass” where each $\text{CH}_2$ unit is defined as 14.000 Da instead of its exact mass of 14.01565 Da. The difference between the exact mass and Kendrick mass is then assessed as the Kendrick Mass Defect (KMD, where KMD = exact mass- Kendrick Mass). The KMD will be constant for compounds within a series that differ only by the number of -$\text{CH}_2$- groups. The elemental formula of the lowest mass member within a series is determined, usually with a high degree of confidence, thereby determining the molecular formulas for all other members of the series. Analysis of mass data from a number of sites shows that most masses can be grouped into pseudo homologous series with mass differences of 14.0156 ($\text{CH}_2$), 2.0157 ($\text{H}_2$; double bond series), and 0.0364 (replacement of $\text{CH}_4$ with oxygen).

Once elemental formulae have been determined, the data are reduced to their elemental H/C and O/C ratios and visualized in a van Krevelen plot (Fig. 2.12). In a van Krevelen plot, the H/C and O/C ratio of each ion can be compared to the elemental ratios.
of likely biochemical precursors, to other samples grouped according to sample location or type, or queried with respect to differences in molecular weight, ionization mode and other features that provide information on DOM composition and cycling. Figure 2.12 shows a van Krevelen plot of FT-ICR-MS from Station ALOHA of the Hawaii Ocean Time-series (HOT). The sample is typical of many open ocean datasets, with the majority of compounds falling within a H/C range of 0.5-1.7 and an O/C ratio of 0.2-0.8 (Kujawinski and Behn, 2006; Koprivnjak et al., 2009; Kujawinski et al., 2009; Gonsior et al., 2011; Hertkorn et al., 2012). Surprisingly, this range of H/C and O/C falls outside the elemental ratios of most common proteins, carbohydrates, and lipids (Fig. 2.12), indicating either SPE-DOM has been extensively reworked and altered from its biochemical precursors, or that SPE-DOM mass spectra target a fraction of carbon that is not abundant in cells. Intensity weighted elemental data from a number of sites shows a narrower range of average values of between 1.3-1.4 for H/C and between 0.3-0.4 for O/C. Ratios may vary somewhat between different studies due to differences in sample handing or instrumental biases. Irrespective of these differences, all studies so far report fairly similar arrays of chemical formulae for SPE-DOM.

In one of the first reports of FTICR-MS data of SPE-DOM in the open ocean, Dittmar and Koch (2006) recognized a cluster of 244 ions with very low H/C and O/C ratios (0.5-0.9 and 0.1-0.25, respectively) and a high number of double bond equivalents (DBE; double bond equivalents, number or double bonds or rings). The high number of DBE and relatively low molecular weights (428-530 Da) that characterize these ions narrowly restricts the structural possibilities. Using conservative estimates of the number of DBE that can be assigned to oxygen, they postulated the low H/C, O/C group represents a class of condensed polycyclic aromatic compounds with 5-8 rings along with different degrees of alkyl substitution and oxygen functionality. As there are no known biogenic precursors for compounds of this type, these compounds most likely originate from thermogenic processes such as terrestrial biomass burning, fossil fuel combustion, and/or reactions of organic matter in hydrothermal systems (Dittmar and Koch, 2006; Dittmar and Paeng, 2009). The discovery of thermogenic polycyclic aromatic compounds in DOM was a significant contribution to our understanding of DOM sources and cycling. Although the potential for a contribution of thermogenic black carbon to
DOM had been recognized earlier from radiocarbon measurements, it was never characterized on the molecular level until the high resolution and specificity of FT-ICRMS was brought to bear on DOM characterization (Masiello and Druffel, 1998).

To quantify polycyclic aromatic carboxylic acids to DOM, an oxidative protocol designed to measure black carbon in soils was adapted to SPE-DOM (Dittmar, 2008). In the method, concentrated nitric acid at elevated temperature (170°C) oxidizes polycyclic aromatic compounds in SPE-DOM to benzene polycarboxylic acids (BPCAs; Fig 2.13), which are separated by HPLC and quantified by on-line spectroscopic detection.

Analyses of marine SPE-DOM yields a suite of substituted BPCAs with a relatively high fraction of benzene penta- and hexacarboxylic acids consistent with highly condensed polycyclic aromatic precursors. The distribution of BPCA isomers varied little between samples, and it is inferred that molecular structures for SPE-DOM are therefore similar, irrespective of sample depth or location. BPCAs represent 1-3% of SPE-DOM.

Subsequent work used the method to quantify total polycyclic aromatic compounds (PCAs) in a meridional section across the Southern Ocean at 30°E (Dittmar and Paeng, 2009). PCA concentrations ranged from 0.6-0.8 µM carbon, or about 1-2% of total DOC. Both high and low values were measured in surface waters, implying inputs and removal of PCAs on relatively short timescales. Due to the remoteness of the Southern Ocean from significant river inputs, atmospheric deposition and subsequent photochemical oxidation were considered to be the most likely sources and sinks for PCAs in this region.

Radiocarbon in BPCAs from HMWDOM recovered by ultrafiltration was highly depleted (-880 to -918 ‰), with ages ranging from 17,000 to 20,100 ybp (Ziolkowski and Druffel, 2010). Concentrations were lower (90-330 nM) than reported by Dittmar and Peang (2009), probably due to lower recoveries of hydrophobic LMWDOM by the ultrafiltration method. However, the results overall support the idea that PCAs represent a refractory DOM fraction. Further studies designed to determine PCA structures are needed to better describe the sources, cycling, and sinks of PCAs.

Studies that compare FT-ICRMS spectra on a simple presence/absence basis of individual ions report a remarkable uniformity in a large fraction of ions from samples collected from different locations and different depths in the ocean. Koch et al. (2005) report that of the 1580 chemical formulae recognized in a suite of samples collected in
the Weddell Sea, ∼30% were present in all samples. Only two formulae were found exclusively in surface waters, and only 79 formulae were found only in deep (>3500 m) water samples. Kujawinski et al. (2009) used statistical analysis of DOM collected in the surface, deep and terrestrial-influenced coastal waters of the US Mid Atlantic Bight, also finding only a very small number of indicator species among > 1000 formulae that could be attributed to exclusively surface marine (32 formulae) or exclusively terrestrial impacted (20 formulae) DOM. Finally, Flerus et al. (2012) found 54% of all masses were present in 90% of 137 samples collected in the eastern Atlantic Ocean, while 74% of mass were present in at least 100 samples. The uniformity in DOM molecular formulae was attributed to refractory DOM with common compositional features that persists over several millennia and is well mixed throughout the entire water column.

ESI-MS of uncharacterized, complex mixtures like SPE-DOM is inherently qualitative. Ionization is selective and subject to matrix effects that are not fully understood. Positive and negative modes yield highly overlapping but different sets of ions due to different ionization efficiencies for carboxylic acids and other functional groups. Only molecules that ionize are registered by the mass detector, and compound classes with low ionization efficiencies (e.g., carbohydrates) are therefore under-represented or can be absent from the mass spectrum (Hertkorn et al., 2006; 2012). However, with careful control of sample processing and analysis, recent studies are beginning to compare the intensity-weighted distribution of molecular formula within sample sets to identify spatial and temporal changes in SPE-DOM composition. Using this approach, Hertkorn et al. (2012) noted an increase in oxygen and a decrease in carbon content between surface (5 m; 36% O; 50-52% C) and deep water (5446 m; 42% O; 47% C) samples, which was attributed to progressive oxidation of DOM with age/depth, in agreement with NMR data that also shows a progressive increase in relative COOH and C=O % carbon with depth. However, Flerus et al. (2012) did not observe a shift in O/C ratios between surface and deep water samples collected on the same cruise. Hertkorn et al. (2012) noted that the abundance of formulae with CHO and CHNO increased with depth relative to formulae containing sulfur (CHOS and CHNOS), in contrast to Kujawinski et al. (2009) who reported an increase in both the number and intensity-weighted number of sulfur containing compounds between the surface (0 m)
and deep (1000 m) Atlantic Ocean. Flerus et al. (2012) used the intensity-weighted approach to distinguish masses that were relatively more abundant in surface waters from masses that were uniformly present (when corrected for total DOC) throughout the water column. They found that masses enriched in surface waters had a lower average mass (300 vs 441 Da), and a lower range in DBE (2-11 vs 7-14), again supported by NMR data suggesting a higher average degree of branching in deep water DOM (Flerus et al., 2012). The intensity-averaged molecular weight increased from 411 Da in surface waters to 417 Da in deep waters, suggesting at best a very small increase in molecular weight with water mass age (Hertkorn et al., 2012; Flerus et al., 2012).

2.4 Links between DOM composition and cycling

The current paradigm of marine DOM cycling draws from a synthesis of rate measurements that span timescales from a few hours to several thousand years (Hansell, 2013; Carlson and Hansell chapter). On very short timescales, marine microbes produce and consume DOM that is “labile” or “reactive”. Annually, a large flux of carbon passes through labile DOC, but at any given moment labile DOC represents only a small fraction (< 0.2 Pg) of the global DOC inventory. Over longer timescales of months to years, excess microbial carbon production and inputs of terrestrial carbon from atmospheric deposition, rivers, and groundwater leads to net accumulation of “semi-labile” or “semi-reactive” DOM in and immediately below the euphotic zone. Net accumulation of semi-labile DOM in the upper water column, and net removal in the mesopelagic ocean, give DOC profiles in temperate and tropical latitudes their characteristic shape, with high values in the surface and lower values at depth. Globally, semi-labile DOM contributes ~20 GT C, or 3% of the marine DOC reservoir (Hansell et al., 2009; 2012). The annual flux of carbon through semi-labile DOM cannot be measured directly, and estimates of residence time span at least an order of magnitude, from a few months (the seasonal accumulation and export of semi-labile DOM from the euphotic zone) to several years (radiocarbon measurements of semi-labile DOM carbohydrates)(Repeta and Aluwihare, 2006; Hansell et al., 2009; 2012). This broad range of rates reflects differences in local production and consumption, differences in how rates are measured, how semi-labile DOM is defined, and the fraction of semi-labile
DOM that is tracked. Lability changes with nutrient conditions, temperature, light, and consumer community structure.

In the deep ocean, DOC values are more stable and decrease only slowly along the path of abyssal circulation. (Hansell et al., 2012) Radiocarbon measurements show deep sea DOC to be several thousand years old (Williams and Druffel, 1987; Druffel et al., 1992). Assuming production and removal of deep sea DOM is a first order process, the turnover time of refractory DOC is ~16,000-30,000 years. The old radiocarbon age and the very slow net removal of deep sea DOC imply that most marine DOM is “non-reactive” or “refractory”. The distinction between labile, semi-labile, and refractory DOM is purely operational. Studies define these terms differently, depending on the nature of the study and the preferences of the investigators. Other nomenclatures exist in the literature which make finer classifications of DOM reactivity (Hansell et al., 2012; Hansell, 2012). However DOM lability is defined, it is inferred from the broad range of timescales over which DOM cycles that composition and lability are linked in some way. One goal of DOM chemical characterization is to identify characteristic features of labile, semi-labile and refractory DOM and understand how composition is linked to cycling.

2.4.1 Composition and the cycling of labile DOM

Grazing, viral lysis of infected cells, and a host of routine cellular physiological processes act to release labile DOM into seawater (Carlson and Hansell chapter). Labile DOM is also produced from the photochemical oxidation of refractory organic matter as it upwells from the deep ocean into the euphotic zone (Kieber et al., 1990; Mopper et al., 1991). Photochemical degradation products include low molecular weight organic acids, aldehydes and ketones that are readily assimilated by marine microbes (Mopper and Stahovec, 1986). The annual carbon flux through labile DOM is measured as heterotrophic bacterial production, under the assumptions of steady state production and consumption and that microheterotrophs are the dominant sink for labile DOM. Measurements of bacterial production are imprecise and do not include organic matter consumed through light driven consumption by photoautotrophs and photochemical oxidation to CO$_2$, or losses due to the adsorption of labile DOM onto sinking particles. Current estimates of bacterial production show that annually, some ~20% of global
primary production is released as labile DOM that is subsequently consumed through a
“microbial loop” in which DOM is either respired or fixed again into microbial biomass
(Fig. 2.14). While most studies have focused on the consumption of DOM by
heterotrophic bacteria and more recently arche, there is abundant evidence that
photoautotrophs also assimilate simple organic acids and other low molecular weight
organic compounds. Some photoautotrophs are auxotrophic for essential vitamins, and
some use siderophores produced by heterotrophic bacteria to supplement their
requirements for iron (Vraspir and Butler, 2009; Helliwell et al., 2011). The exchange of
labile DOM between auto- and heterotrophic microbes therefore limits and shapes
microbial production and marine biogeochemical cycling in a very fundamental way.

Labile DOM constituents include easily characterized, low molecular weight
biochemicals (simple amino acids, sugars, organic acids, ATP, vitamins, etc.) that are
readily assimilated by marine microbes and simple biopolymers (proteins, unbranched
homopolysaccharides, etc.) that can easily be hydrolyzed by extracellular hydrolytic
enzymes. The high demand for labile organic substrates, nutrients, and metals by the
microbial community keeps steady state concentrations of labile DOM constituents at
nanomolar levels (Fuhrman, 1987; Skoog and Benner, 1997; Kaiser and Benner, 2009).

As discussed in section 2.3.2, concentrations of dissolved free amino acids and THAA
range from a few nM (for free amino acids) in the deep ocean, to 100s of nM (for THAA)
in the surface open ocean. Radiocarbon tracer experiments show free amino acids are
rapidly metabolized by heterotrophic bacteria, and the uptake and incorporation of the
tritiated amino acid leucine is commonly used as a measure of bacterial production
(Fuhrman, 1987). The lability of specific proteins and bacterial peptides that constitute
THAA is unknown, but spatial and temporal changes in THAA concentrations along with
experiments tracking extracellular peptidase activity by marine phytoplankton and
bacteria show a significant fraction of THAA are metabolized over short time scales
(Amon et al., 2001). Dissolved simple sugars likewise occur at low nM concentrations in
the euphotic zone and are taken up quickly by bacteria (Rich et al., 1996). Total
hydrolyzable neutral and amino sugars concentrations range from > 600 nM in open
ocean surface waters to 30-60 nM at depth. It is unclear what fraction of DOM
polysaccharides are labile. Extracellular, dissolved polysaccharides are universally
produced by marine algae and cyanobacteria when grown in laboratory pure culture (Aluwihare et al., 1997; Biddanda and Benner, 1997; Aluwihare and Repeta, 1999; Meon and Kirchman, 2001). A number of studies have explored carbohydrate degradation by inoculating filtered, spent culture media with heterotrophic bacteria and following simple sugar composition or DOM spectral characteristics over time. In these experiments homopolysaccharides are rapidly degraded by heterotrophic bacteria, but degradation is always selective; the composition of neutral sugars invariably changes over the course of the degradation experiments (Aluwihare et al., 1997; Meon and Kirchman, 2001). In the field, experiments using fluorescently labeled model compounds also show large changes in polysaccharide hydrolysis rates and substrate selectivity with location (Arnosti et al., 2011; Arnosti, 2011). Labile polysaccharides are readily produced and consumed during upper ocean carbon cycling, but steady state concentrations remain low as with other labile components of DOM due to tight coupling between production and consumption by marine microbes. Although proteins and polysaccharides contribute the majority of carbon and nitrogen in living microbial biomass and marine particulate organic matter, microbes produce an enormous diversity of organic compounds as part to their metabolic systems. Untargeted characterization of DOM shows hundreds to thousands of different compounds can be recovered by SPE from pure cultures of marine photoautotrophs. Initial results suggest chemical diversity may track phylogenetic diversity, such that the genomic variability in marine microbial communities may be matched by a similar chemical variability in labile DOM composition. Some labile DOM constituents may be generally available to the heterotrophic community, but others induce highly specific growth responses. For example, studies using defined radiolabeled substrates or labeled DOM produced in pure cultures of specific marine photoautotrophs show selective uptake by defined subclasses of marine heterotrophs (Sarmiento and Gasol, 2012; Cottrell and Kirchman, 2000). DOM released by marine heterotrophs likewise enhances the growth of only certain marine autotrophs. A recent review highlights the interaction of marine bacteria and diatoms, while other studies have documented enhanced growth of the marine cyanobacteria Prochlorococcus in the presence of some heterotrophic bacteria (Sher et al., 2011; Amin
Prochlorococcus-heterotrophic mutualism is highly strain specific, some strains of bacteria enhance growth of only some strains of Prochlorococcus.

In the examples above, the DOM constituents that lead to enhanced growth have not been identified, but B vitamins and siderophores are two examples of labile DOM nutrients that are known to impact microbial community production and diversity.

Approximately half of all marine microalgae, including many harmful algal bloom forming species, as well as many marine bacterioplankton, are auxotrophic for vitamins B$_1$ and/or B$_{12}$ (Tang et al., 2010; Helliwell et al., 2011). B-vitamins co-limit and shape marine microbial production, so knowledge of the distribution of B vitamins is important to assessments of nutrient limitation (Tang et al., 2010; Bertrand et al., 2011; Bertrand et al., 2012). Early measurements of B vitamin distributions relied on bioassays, but recently, direct analysis of B vitamins by mass spectrometry has been reported.

Concentrations of B$_{12}$ reach low nM levels in some coastal waters, but fall to < 1 pM (detection limit) at open ocean sites due to low bacterial production and high uptake by B-vitamin auxotrophs (Sañudo-Wilhelmy et al., 2012).

Over 99% of dissolved iron in the ocean is complexed to organic ligands of unknown composition (Gledhill and van den Berg, 1994; Rue and Bruland, 1997; Boye et al., 2001; Gledhill and Buck, 2012). In many remote areas of the equatorial Pacific, northwest Pacific, and Southern Oceans iron concentrations fall to < 100pM, creating enormous selective pressures for microbes to develop efficient Fe uptake and utilization strategies (Martin and Fitzwater, 1988; Martin et al., 1991; De Baar et al., 2005; Jickells et al., 2005; Follows et al., 2007; Dutkiewicz et al., 2009; Bragg et al., 2010; Barton et al., 2010; Boyd and Ellwood, 2010; Dutkiewicz et al., 2012; Miethke, 2013). Microbes that can access iron-organic ligand complexes have a distinct competitive advantage in iron-depleted regions of the ocean. One mechanism by which microbes can acquire iron is through the siderophores, strong iron-organic complexes that facilitate iron transport and uptake across the cell membrane (Vraspir and Butler, 2009). Like B vitamins, dissolved iron-organic complexes occur at extremely low concentrations that are difficult to measure. However, recent reports describe a number of siderophore and siderophore-like compounds in seawater at pM concentrations, and advances in analytical technologies should facilitate future measurements (Mawji et al., 2008; Mawji et al., 2011; Velasquez...
et al., 2011; Gledhill and Buck, 2012; Boiteau et al., 2013). The distribution and cycling of vitamins and trace metal organic complexes are two developing areas of research that target specific, labile DOM constituents with measurable impacts on marine microbial communities. Studies of untargeted labile DOM composition and microbial interactions suggest that as yet unidentified labile DOM constituents may also influence an array of metabolic processes from cell-cell signaling (e.g., quorum sensing) to nutrient uptake. Understanding labile DOM composition is key to understanding many interactions between microbial communities, as well as many of the processes that limit and shape the microbial community production and diversity.

2.4.2 Composition and the cycling of semi-labile DOM

Traditionally, semi-reactive DOM has been defined as the fraction of DOC that is present in surface waters but that is not at depths > 1000m, the point where DOC concentration values begin to stabilize. This definition arises from two principal observations; 1) DOC concentrations in temperate and tropical surface waters are elevated relative to deep ocean values, and 2) radiocarbon models of DOC age use a mix of modern and aged components to explain upper ocean $\Delta^{14}C$-DOC values (Druffel et al., 1992; Hansell et al., 2009; 2012). However, semi-labile DOM can also be defined on the basis of turnover time relative to labile and refractory DOM fractions. Viewed this way, semi-labile DOM is any organic matter that cycles on seasonal to decadal timescales. This definition includes DOM in the deep ocean (> 1000 m) that is in quasi-steady state with its production (via dissolution of sinking particles, in situ microbial production, etc.) and removal (Repeta and Aluwihare, 2006; Orellana and Hansell, 2012). The timescales over which semi-labile DOM cycles are too long to be captured experimentally in laboratory or field incubation experiments without significant bottle effects, but too short to be captured by natural abundance radiocarbon measurements. Linking semi-labile DOM constituents to cycling pathways is therefore difficult, and the processes that result in the net accumulation of semi-labile DOM in the surface ocean, and net degradation at depth, remain largely unexplained.

Marine algae and bacteria synthesize semi-labile DOM polysaccharides and proteins, which in laboratory degradation experiments can persist for several years.
Heterologous carbohydrate amendments to seawater simulate rapid increases in bacterial carbon utilization by a succession of bacterial taxa, suggesting resource partitioning of semi-labile DOM between different species of bacteria (McCarren et al., 2010). Coordinated, cooperative interactions by a variety of different bacteria may be necessary to degrade the complex, highly branched acylated polysaccharide that represents a large fraction of semi-labile DOM. The carbon respired by short-term incubation experiments of this type represents only a very small fraction (typically < 1%) of the carbon amendment, and the link between short term incubation results and semi-labile DOM degradation in situ is not clear. Longer incubations of nutrient amended surface waters show greater losses of DOC, but still fall far short of consuming a major fraction of the semi-labile DOC reservoir. In a series of experiments, Carlson and colleagues measured DOC respiration in a matrix of surface and 250 m water treatments that had been 0.2 µm filtered, amended with nutrients, and inoculated with either unfiltered (containing bacteria) surface or 250 m water (Carlson et al., 2002; 2004). No significant DOC respiration and only a slight increase in bacteria cell numbers were observed in surface water inoculated with surface water bacteria. However, significant respiration (up to 10% of total DOC) and 3x increases in bacterial cell numbers were measured in treatments of surface water with 250 m bacteria. The studies of Carlson and colleagues show that the bacterial community in surface waters has a limited capability to degrade polysaccharides and other components that contribute to semi-labile DOM, while mesopelagic bacterial communities have the requisite metabolic pathways to metabolize at least a fraction (up to 20%) of semi-labile DOM. The results of microbial degradation experiments have yet to be coupled to studies of semi-labile DOM composition. Characterizing the small amounts of semi-labile DOM that is consumed in these incubation will prove to be challenging, and developing improved experimental approaches to simulating semi-labile DOM degradation will be key to future progress in this area.

Metabolism of DOM by heterotrophic bacteria is selective and/or leads to subtle transformations of semi-labile DOM. A growing number of studies have measured differences in DOM composition between coastal and open ocean DOM or between surface and deep water DOM and tied these differences to the “apparent diagenetic state”
or potential reactivity of semi-labile DOM. Goldberg and colleagues measured seasonal changes in DOC, total carbohydrates (THCO) and dissolved combined neutral sugars (DCNS) as DOM accumulates in surface waters at the Bermuda Atlantic Time-series Study site (BATS; 31° 40’N, 64°10’W) and spatially in surface waters across the North Atlantic subtropical gyre between 7-43°N (Goldberg et al., 2009, 2010). Values for DOC, THCO, and DCNS were highest during summertime stratification of the water column and in the mid-gyre (~ 24°-27°N). Recently accumulated DOM had higher carbohydrate yields (%THCO) and higher mol% galactose and mannose+xylose than DOM at depth, which was characterized by higher mol % glucose. As semi-labile DOM is processed by bacteria, it becomes less amenable to chemical characterization and the distribution of DCNS changes (Amon and Benner, 2003; Goldberg et al., 2009; 2010).

Patterns of amino acid and hexosamine distributions in particulate organic matter and sediments are also strongly imprinted with the combined effects of degradation and transformation. By comparing the distribution of amino acids in surface and subsurface sediments, and in fresh and recycled particulate organic matter, Dauwe proposed a degradation index that relates the composition of organic matter to lability (Dauwe et al., 1999). The degradation index assumes steady state deposition over time, so that the distribution of amino acids/hexosamines in contemporary surface sediments or fresh particulate matter is a good representation of the amino acid distribution of older sediments or more processed particulate organic matter at the time of synthesis. This concept has been applied to studies of DOM cycling to provide a measure of DOM degradation. For semi-labile DOM, it is assumed that the distribution of amino acids/hexosamines in coastal surface waters is indicative of the initial product, which is transformed as waters move offshore or into the deep sea. Yamashita and Tanoue (2003) observed decreases in the %THAA and increases in the relative amount of glycine and alanine in a suite of samples from near shore (fresh DOM) to offshore and greater depths (more recycled DOM), concluding that degradation had a major impact on the quantity and quality of THAA (Yamashita and Tanoue, 2003b). Likewise, Davis et al. (2009) measured the %THAA and degradation index in a suite of (organic matter) amended and unamended Arctic Ocean seawater samples, finding that %THAA changed rapidly in the days immediately following amendment, but stabilized and changed only slightly.
thereafter. These changes were accompanied by minor changes in the amino acid degradation index, which varied in the upper 200 m of the water column but was invariant below this depth. These and other studies show that freshly produced organic matter has a relatively higher %THAA, which decreases on timescales of weeks to months, while the ratio of indicator amino acids can be used to track changes occurring over timescales of months to years.

Hubberten et al. (1994; 1995) used hydrophobic organic resins (XAD-2; a cross-linked polystyrene) to extract amino acids from the Greenland Sea and Southern Ocean. Seawater was filtered (Whatman GF/F), acidified to pH = 2 with hydrochloric acid, then passed through the XAD column. (Hubberten et al., 1994; 1995) Amino acids incorporated into hydrophobic substances are adsorbed onto the column, then recovered by sequentially rinsing with 0.2N NaOH and methanol. Each fraction was analyzed for amino acids. In all samples from < 1000m, TDAA concentrations always exceeded amino acids in the XAD fractions (XAD-AA; Fig. 2.7). However, the difference, the amount by which [TDAA] exceeded [XAD-AA], was greatest for surface waters and diminished with depth. Hubberten noted a linear correlation between chlorophyll-a [Chl-a] and THAA in surface waters. Extrapolation to [Chl-a]= 0 yielded a positive intercept for [THAA] that was similar to values of XAD-AA. For depths below 100m, concentrations of THAA and XAD amino acids were approximately equal. Both observations led Hubberten et al. to suggest that there are at least two fractions of hydrolysable amino acids in seawater, a labile fraction that is largely hydrophilic and present in surface waters and a semi-labile or refractory component of about 150-250 nM present throughout the water column. The distribution of amino acids in the XAD fraction was characterized and found to be similar to THAA, so no distinction of the two fractions could be made by amino acid distribution. No further characterization of XAD-AA was made, however proteins, peptides, and bacterial cell wall material have some hydrophobic character, and can be retained on hydrophobic resins under low pH conditions. The amino acids recovered by Hubberten and colleagues could have included amino acids bound in humic substances, proteins/cell wall biopolymers and their degradation products. The decrease in the ratio THAA/XAD-AA with depth suggests some change in the macromolecular form of amino acids, a point corroborated by amino
acid fluorescence data, which shows only tryptophan-like fluorophores in surface waters, but additional tyrosine-like fluorophores in deep water samples. (Yamashita and Tanoue, 2004)

Finally, DOM in filtered seawater spontaneously assembles into polymer gels, stable three-dimensional networks of DOM macromolecules that grow until they reach a size and density that allows them to sink or adhere to sinking particles; Orellana chapter. Assembly is rapid and reversible, and gels are one mechanism by which semi-labile DOM can be removed from the surface ocean. As discussed in section 2.3.3, RuBisCo is produced in the euphotic zone, transported to the deep ocean by sinking particles, and released by DOM-POM exchange. In regions of high particle flux such as the equatorial Pacific, RuBisCo can be detected well into the deep ocean where it persists as semi-labile DOM. Gels stain positively for carbohydrates, proteins, and lipids and these other constituents of polymer gels are likely carried into the deep ocean as well.

Peptidases, proteases, and ATPase enzymes have been identified in polymer gels through proteomic analyses (Orellana et al., 2007). No similar analyses of polymer gel carbohydrates have been reported, but bubble formation and collapse concentrates polymer gels into sea surface foam, and chemical analyses of natural sea foams have been made (Orellana et al., 2011). NMR spectra of natural sea surface foam and foam produced after the microbial degradation of spent algal culture media have spectral features characteristic of acylated polysaccharides that accumulate in surface water as semi-labile DOM carbohydrate. Combined neutral sugar distributions of semi-labile DOM and foam are also similar (Gogou and Repeta, 2010). $^1$H NMR and $^{13}$C NMR spectra of DOM collected by ultrafiltration, SPE, and ED/RO all show resonances assigned to semi-labile DOM carbohydrate as deep as 5200 m, but further chemical and isotopic characterization is needed to link surface and deep reservoirs of semi-labile DOM (Benner et al., 1992; Sannigrahi et al., 2005; Quan and Repeta, 2007; Helms et al., 2013).

2.4.3 Composition and the cycling of refractory DOM

By definition, refractory DOM does not degrade via the typical microbial and chemical processes that recycle labile and semi-labile DOM. Understanding why
refractory DOM is so unique is therefore critical to a comprehensive description of the marine carbon cycle. The chemical composition of SPE-DOM is taken to be representative of refractory DOM. As discussed in section 2.3.3, SPE-DOM is a very complex mixture of at least several thousand different components with elemental H/C and O/C ratios that lie outside the range of common lipids, proteins and carbohydrates (Fig. 2.12). SPE-DOM is therefore not thought to have a direct biological source, but the case for direct biological production and selective preservation of refractory organic matter has been made for sediments, and given the very small flux of carbon (< 0.04 Pg year⁻¹) needed to support the radiocarbon age of refractory DOM, and the large annual flux of carbon through marine primary and secondary production, direct synthesis of refractory DOM cannot be discounted. The complex molecular features of SPE-DOM can be interpreted as the result of transformation of lipids, proteins and carbohydrates that become scrambled in such a way as to make them difficult for marine microbes to assimilate or metabolize. Refractory DOM is characterized by a high proportion of carboxylate and fused alicyclic ring carbon that shares structural characteristics with polycyclic lipids (Hertkorn et al., 2006). The mechanisms that could convert simple lipids to refractory DOM are a matter of conjecture, and more work is needed to assess both the feasibility of the formation pathway and its impact on microbial metabolism.

Although refractory DOM is not readily metabolized by marine microbes, a number of sinks that rely on physical/chemical processes have been identified. In a series of elegant experiments Mopper and co-workers detailed the photochemical oxidation of deep sea colored dissolved organic matter on exposure to light (Mopper and Stahovec, 1986; Kieber et al., 1990; Mopper et al., 1991). Photooxidation products include low molecular weight, labile, organic acids, aldehydes and ketones that are rapidly consumed by bacterial heterotrophs. Photochemical oxidation also decreases the adsorptive and fluorescent properties of DOM through the transformation or removal of unsaturated functional groups (Kieber et al., 1990; Weishaar et al., 2003). Subsequent studies have confirmed and expanded these observations and a photochemical sink and transformation pathway for otherwise refractory DOM is now firmly established. However, photochemistry is confined to the upper portion of the euphotic zone and the amount of carbon potentially mineralized by this pathway is very limited. Photochemical oxidation
of DOM probably leads to significant isotopic fractionation between products, but the isotopic value of DOM is similar to particulate organic matter, and it is unlikely that photochemical oxidation is the primary sink for refractory DOM.

Globally, the largest sinks for refractory DOM are in the deep sea (Hansell et al., 2009, 2012; Hansell, 2012; Hansell and Carlson, 2013). From an analysis of salinity and high precision DOC values in the deep Pacific Ocean, Hansell and Carlson (2013) were able to identify two regions of refractory DOC removal within the basin, a deep sink in the far North Pacific, and a mid depth sink in the tropical South Pacific. They estimated that these two sinks remove 7-29% of the 43 TG refractory DOC introduced into the deep global ocean by overturning circulation. Two processes, removal during hydrothermal circulation, and adsorption onto sinking particles, have been identified as sinks for deep sea refractory DOM. Hydrothermal fluids exiting permeable ocean crust at mid-ocean ridge crests and unsedimented/thinly sedimented ridge flanks have DOC concentrations that are lower by 20-25 µM than deep seawater (36-39 µM; typical vent fluid DOC concentrations are 11-19 µM; Lang et al., 2006; Lin et al., 2012). Global DOC losses in hydrothermal systems are low, < 0.0002 Pg C year⁻¹, but measurements are few and this may be an underestimate. The composition and lability of the ~20 µM DOC that exits with hydrothermal fluids has not been studied, but radiocarbon values of the HMWDOC fraction range from -772‰ to -835‰, significantly depleted relative to overlying seawater (McCarthy et al., 2010). Other processes such as microbial oxidation and removal by self-assembling organic gels may be active and need further study.

Radiocarbon values of suspended particulate matter decrease by almost 100‰ with depth in the ocean (Druffel et al., 1992;1998). Even at the low Stokes settling velocities calculated for µm sized particles (~ 1m day⁻¹), suspended particulate matter settles quickly enough that no appreciable gradient in radiocarbon values due to aging should be measureable. To explain the gradient in suspended POC radiocarbon values, Druffel and colleagues postulated that radiocarbon depleted, refractory DOM is adsorbed onto suspended particles. Adsorption forms the basis for SPE extraction of refractory DOM, and it would be surprising if a similar process does not occur in the water column that is permeated with mineral surfaces and organic particles. A number of factors including complex POM dynamics in the deep ocean, resuspension and advection of fine-
grained sediments with depleted radiocarbon values from continental margins, and
adsorption of refractory DOM onto filters used to collect particles, potentially influence
the DOM-POM adsorption model. However, based on Druffel’s data, Hansell calculated
that adsorption could remove ~0.05 Pg C yr\(^{-1}\) from the deep ocean, enough to sustain the
observed deep-sea gradient in DOC concentration (Hansell et al., 2009). No analyses of
particle-adsorbed, refractory DOM have been attempted. Given the recent progress in
SPE-DOM characterization, and the ~500 ‰ separation in radiocarbon values for newly
synthesized POM (+ 50 ‰) and refractory DOM (-400‰), it might be feasible to test the
adsorption/removal hypothesis through a combination of structural and isotopic analyses.

2.5 Future research

Over the past decade, our understanding of DOM composition and cycling has
advanced in a number of areas. Reverse osmosis/electrodialysis now provides a larger
fraction of DOM for study than either ultrafiltration or solid phase extraction alone. High
field, multi-dimensional NMR and high resolution MS are providing unprecedented
details of DOM composition, and new techniques of data integration and visualization
such as correlation spectroscopy are allowing marine chemists to couple different spectral
datasets into a more integrated picture of DOM composition. This wealth of new
information presents both opportunities and challenges to the DOM community. One
challenge is to provide community-wide access to the very large amount of mass and
NMR spectral data available through open-access databases. Presently, such data are
often only summarized or provided as supplemental materials in published reports. A
database of spectral information that includes details of the methodologies needs to be
made available for interrogation by the wider scientific community if the full benefits of
advanced methods of characterization are to be realized.

Much more effort needs to be placed on the validation of results by bringing
different analyses to bear on common samples or at common study sites. It is unlikely
and perhaps undesirable for sampling methods to converge at this time on one technique
that provides the highest recovery of DOM. Each method has particular benefits of cost,
speed, throughput, and selectivity for different fractions of DOM. A better understanding
of the sampling overlap between techniques would however help to interpret similarities
and differences in composition between SPE, ultrafiltration, and ED/RO samples. Most current studies focus on one or two methods of spectral characterization that are within the scope of expertise of particular analysts. Collaborative studies that integrate data from a suite of different spectral (MS, NMR, IR, etc.), chromatographic (HPLC, GC, etc.), elemental and isotopic ($\delta^{13}$C, $\Delta^{14}$C, etc.) analyses would allow for more robust interpretation of complex datasets. For specific components of DOM that have biomarker or tracer potential (thermogenic DOM, APS), more comprehensive isotopic and structural characterization, along with better methods for isolation and purification, would substantially improve existing models of DOM cycling.

Advances in technology have made in-depth studies of labile DOM composition and cycling feasible. Automated, high resolution multi-dimensional liquid and gas chromatographic systems coupled with high resolution MS and high throughput microcapillary NMR offer the potential to better characterize the exometabolome of labile DOM. A large number of model photoautotrophs and heterotrophs representing major ecotypes of marine cyanobacteria, algae, and bacterioplankton are now available as pure cultures with sequenced genomes. Labile DOM characterization of these cultures at different growth stages would provide valuable links between genomic and exometabolomic composition, as well as microbial dynamics in natural systems. Paired co-cultures studies further demonstrate that synergistic interactions between photoautotrophs and heterotrophs are probably common in the ocean. Little is known about the organic nutrients and substrates that stimulate these associations it is a fruitful area for further study, particularly when paired with transcriptomic and proteomic datasets. There are only a few reports of the distribution of vitamins and other bioessential organic compounds (siderophores, quorum sensing compounds, etc.) in seawater. These compounds have a demonstrable impact on microbial ecosystems, but their production and uptake are only poorly understood. Finally, labile DOM composition needs to be linked to genomic studies already underway at time-series study sites and in global surveys of microbial populations and water column chemical properties. Labile DOM composition might provide an important new input to global models of microbial metabolism, diversity, and community structure.
$^{13}$C NMR spectra of DOM collected by RO/ED provides the most comprehensive view of semi-labile and refractory DOM composition currently available (Fig. 2.15). These spectra show the presence of two major fractions, a semi-labile HMWDOM fraction with acylated polysaccharide (APS) as the major component, and a refractory fraction with carboxy-rich aliphatic matter (CRAM) as the major component. The NMR spectral, elemental (C/N/P), and partial chemical (monosaccharide composition, linkage pattern) characteristics of APS are known. This level of detail has allowed for the detection of APS throughout the water column, but has not allowed for the identification of major sources of APS, or provided either an explanation for why APS accumulates in the upper ocean or a better estimate of carbon flux through semi-labile HMWDOM. The major sinks for APS have also not been identified. Better characterization of APS using mass spectrometry and 2D NMR is an essential first step in addressing these questions, as are improved chemical techniques that can fully depolymerize APS for comprehensive monosaccharide and linkage analyses. A more detailed description of APS composition should allow for better distinction of APS sources in laboratory culture experiments, would help in the design and implementation of microbial degradation experiments, and experiments monitoring gel polymer formation as putative sinks for APS. Finally, there is evidence from the study of proteins in DOM for the transport of semi-labile DOM into the deep ocean. Determining the inventory of semi-labile DOM in the deep ocean, its radiocarbon value, and its sinks would provide important new insights into deep sea carbon cycling and perhaps bathypelagic microbial ecology.

Understanding refractory DOM may be the most formidable challenge for DOM composition and cycling. High resolution mass and high field NMR spectral analyses have yielded unprecedented details into refractory DOM composition, but the results need to be independently verified by other techniques and made more quantitative. For example, the current view of SPE-DOM composition is of a very complex mixture of low molecular weight carboxyl-rich aliphatic compounds. If correct, these compounds should be amenable to separation and characterization by multi-dimensional gas chromatography-mass spectrometry (GC-MS). 2D GC-MS has already proved to be a valuable tool in the characterization of petroleum, which is a mixture polycyclic alkanes.
and alkenes, some of which incorporate nitrogen and sulfur, similar in many respects to
the proposed composition of refractory DOM.

Better integration of carbon isotope analyses with structural characterization is
also needed. The isotopic composition of thermogenic DOM should provide important
insights into the origin and cycling of polycyclic aromatic compounds in DOM. Only
one such set of measurements have been made, and more are needed (Ziolkowski and
Druffel, 2010). Finally, there are only a few reports of deep sea DOM spectral
characteristics. Given the large gradient in concentration and radiocarbon value through
the deep ocean, it seems likely that aging and removal of DOM will impact DOM
composition. A survey of DOM composition through the deep ocean may offer insights
into the processes that cycle refractory DOM.

There is observational evidence for two deep ocean sinks of refractory DOM,
removal during hydrothermal circulation of seawater through permeable crust and
adsorption onto particles. Both sinks merit further study. The composition of DOM in
vent fluids needs to be determined and compared with the isotopic and chemical
composition of refractory DOM. Preliminary work suggests that hydrothermal
circulation is not a large source or sink of refractory DOM, but further measurements are
needed to verify this inference. Adsorption onto sinking particles has been proposed but
not verified through experiments using molecular level analysis. Ultrafiltration and
ED/RO or SPE allow for the concentration of deep sea POM and DOM that could be
used in to experimentally track adsorption. Given the high level of detail that can now be
achieved using advanced analytical methods, it may also be feasible to isolate compounds
with chemical and isotopic characteristics of refractory DOM from suspended particulate
organic matter. Such studies would help to verify the adsorption hypothesis. Nothing is
known about microbial degradation of refractory DOM.

Finally, little is known about the role of deep sea microbes in DOM removal.
Rates of microbial DOM cycling are thought to be very low and therefore difficult to
simulate in laboratory experiments. However, using more sensitive and precise methods
of DOM characterization that are now available, it may be possible to explore deep sea
microbial/refractory DOM interactions to provide new insights into this potentially
important aspect of the marine carbon cycle.
2.6 Acknowledgements

The author has had the good luck to work with a large number of talented scientists over many years whose enthusiasm, ideas, and laboratory work have contributed to this chapter. In particular this chapter has benefited from discussions with Lihini Aluwihare, Carl Johnson, Chritos Panagiotopolous, Mar Nieto Cid, Aleka Gogou, Chiara Santinelli, Robert Chen, Jamie Becker, Rene Boiteau, and Chris Follett. Tracy Quan. Hussain Abdulla and Eiichiro Tanoue provided many helpful comments and suggestions that significantly improved the manuscript. I would also like to thank the NSF Chemical Oceanography and National Science and Technology Center for Microbial Oceanography Research and Education (C-MORE) programs for their generous support over many years, and the Gordon and Betty Moore Foundation for their interest and support of microbial cycling and DOM.
Figure captions

**Figure 2.1** Extraction methods for marine DOM. In solid phase extraction (left) the sample is passed through a column packed with an organic sorbent (polystyrene, octadecylsilyl coated silica gel, etc.). Physisorptive attraction between hydrophobic DOM and the sorbent leads to adsorption of DOM on the column, which in the photograph appears as a brown discoloration (due to the adsorption of colored dissolved organic matter). After the sample has been extracted, hydrophobic DOM is recovered by washing the column with methanol. In ultrafiltration (top, right), seawater is pressurized and passed through a filter with nanometer-sized pores (grey sheet in figure). Organic molecules with hydrodynamic diameters greater than the filter pore size (red ellipses) are retained, while salt, low molecular weight DOM, and water (blue circles) permeate through the filter. Once the sample has been concentrated to < 1-2 L, residual salts are removed by serial dilution with ultra-pure water followed by filtration. In reverse osmosis/electrically assisted dialysis (RO/ED; bottom right) the sample is processed in serial ED and RO steps. In the ED step, the sample is passed through stacks of cation and anion exchange membranes that are under the influence of an applied electric potential and washed by alternating flows of sample (seawater, black arrows) and ultra-pure water (freshwater; grey arrows). Anions (chloride, sulfate, etc.) pass through the positively charged membrane (blue panels) toward the cathode, while cations (sodium, potassium) pass through the negatively charged membranes (green panels) toward the anode. This transfers salts from the sample into the freshwater feed, lowering the salinity of the sample. The sample is then concentrated by reverse osmosis, and processed again by ED to further reduce the salt content of the sample. Figure adapted from http://www.osmo-membrane.de.

**Figure 2.2** Nuclear magnetic resonance spectra of HMWDOM collected by ultrafiltration of surface seawater collected from the North Pacific Ocean using a polysulfone membrane with 1 nm pore size (nominal 1 kDa molecular weight cut-off). (A) $^{13}$C NMR spectra have major peaks from carboxyl (COOH, CON; ~175 ppm), aromatic and olefinic C (broad peak centered at ~ 140 ppm), anomeric (OCO; 110 ppm), O-alkyl C (HC-OH; 70 ppm), methine and substituted methylene C (~35-40 ppm), and two alkyl C peaks from acetamide (26 ppm) and 6-deoxysugars (20 ppm). (B) $^1$H NMR spectra show major peaks from anomeric protons (5.2 ppm), O-alkyl protons (3.5-4.5 ppm), acetamide methyl (2.0 ppm) and 6-deoxy sugar methyl and methylene protons (1.3 ppm). These peaks sit atop a broad baseline between 0.9-4 ppm from substituted methine, methylene, and methyl protons from hydrophobic DOM.

**Figure 2.3** (A) Overlay of $^1$H NMR spectra from North Pacific Ocean HMWDOM before (black trace) and after (red trace) passage through an anion exchange resin. Ultrafiltration concentrates from seawater high molecular weight, acylated polysaccharides (APS) as well as low molecular weight, hydrophobic humic substances. These two chemically distinct fractions of DOM can be separated by solid phase extraction or ion exchange chromatography, providing relatively pure fractions of acylated polysaccharide (APS) and carboxyl-rich aliphatic matter (CRAM). (B) Contour map of synchronous changes in peak intensity for $^{13}$C NMR spectra of HMWDOM.
collected at estuarine and coastal marine sites. Positive correlations between functional
groups appear in red as off-diagonal peaks. Negative correlations appear in blue, with the
intensity of the colors scaled to the intensity of the correlation. Figure courtesy of Drs.
Hussain Abdulla and Patrick Hatcher. Samples were collected using a polysulfone
membrane with a 1 nm pore size.

Figure 2.4 $^{15}$N NMR spectra of HMWDOM (polysulfone membrane with 1 nm pore
size) before (top) and after (bottom) acid hydrolysis. HMWDOM nitrogen occurs
primarily as amide-N (124 ppm) with smaller amounts of amine-N (35 ppm). Major
biopolymers that incorporate amide-N are proteins and N-acetyl-aminosugars. The
contribution of proteins and N-acetyl-aminosugars in HMWDOM can be assessed by
combining $^{15}$N NMR spectral data with measurements of acetic and amino acids released
after acid hydrolysis. Treatment of HMWDOM with acid converts most amide-N into
amine-N. For surface waters, companion chemical analyses show that N-acetyl amino
sugars contribute 40-50% of amide-N while proteins contribute 8-13% of amide-N.
Figure adapted from reference 8.

Figure 2.5 Representative structures of monosaccharides isolated from HMWDOM
hydrolysis products. Treatment of HWMDOM with acid yields a suite of hexoses
(mannose, galactose, glucose), a pentose (arabinose), 6-deoxyhexoses (fucose,
rhamnose), hexosamines (glucosamine, galactosamine), and a large number of
methylated hexoses (3-methyl rhamnose, etc.) as well as some novel deoxysugars (3-
deoxyglucose).

Figure 2.6 2D NMR homonuclear correlation spectroscopy (COSY) of HMWDOM
carbohydrate from the North Pacific Ocean (Fig. 2.2). In this example, strong cross
peaks between H-5 x H-6 (red) and H-2, H-3, and H-4 (blue) show the presence of 6-, 4-, 3-, and 2-deoxysugars in HMWDOM. Additional strong cross peaks are observed
between $\alpha$H-1 x H-2, and $\beta$H-1 x H-2 (violet), showing different stereochemical
geometries within the polysaccharide.

Figure 2.7 Distributions of total hydrolysable amino acids (THAA; left panel), D-amino
acids (center panel) and SPE-amino acids and THAA (right panel). Data in the left and
center panels are from the North Pacific (grey squares; Station ALOHA of the Hawaii
Ocean Time-series (HOT) program) and North Atlantic (open circles; Bermuda Atlantic
Time-series Study (BATS) program) Oceans. SPE-AA (filled triangles) and THAA
(open triangles) data in the right panel are pooled datasets from the high latitude North
Atlantic and Southern Oceans. Concentrations of THAA are high in the euphotic zone
and fall through the mesopelagic zone. Concentrations stabilize in the deep ocean, >1000
m. D-amino acids concentrations are also highest in the euphotic zone, decrease to about
1000m and stabilize thereafter. SPE-AA are thought to represent amino acids and
peptides that are part of refractory DOM, showing much less variation in concentration
with depth. At 2000m, nearly all THAA is recovered as SPE-AA. Data from HOT and
BATS are from reference 49. Data on SPE-AA and THAA in the right panel are from
reference 140.
Figure 2.8 The relative distributions of individual amino acids within dissolved THAA, phytoplankton, suspended particulate matter, and sinking particles. Plankton, suspended, and sinking particle data are from the equatorial Pacific Ocean, while THAA is from Station Aloha, near Hawaii. (Lee et al., 2000; Kaiser and Benner, 2009) Sinking particle data are from long-term deployments of floating sediment traps. Data from long-term deployments of moored traps have higher concentrations of glycine (Gly), similar to THAA in DOM. All sample types have the same suite of amino acids. The major difference between dissolved and particulate amino acid distributions are the higher relative abundance of aspartic acid and glycine in THAA. Amino acids are: Asp (aspartic acid), Glu (glutamic acid), Ser (serine), His (Histidine), Thr (Threonine), Gly (Glycine), Arg (Arginine), Ala (Alanine), Tyr (Tyrosine), Val (Valine), Met (Methionine), Ile (isoleucine), Leu (leucine), Lys (lysine).

Figure 2.9 SDS-PAGE gel of dissolved proteins in seawater visualized by silver staining. Dark bands represent separated proteins. The left and right lanes are standard mixtures of known proteins used to calibrate the molecular separation of the gel. Sample proteins are from surface waters collected from the equator (Station 1) to 60°S (Stations 4, 5) along 120°E, then west across the Southern Ocean to approximately 30°E (stations 6-9). Note the similarity in protein bands between samples. Major protein bands appear in samples at 48 kDa, 37 kDa, and 15 kDa. Figure used with kind permission from Dr. Eiichiro Tanoue.

Figure 2.10 ¹H NMR of SPE-DOM from 900m near Hawaii. NMR spectra of SPE-DOM differs from ¹H NMR spectra of HMWDOM by the absence of carbohydrates. Major resonances include methyl (H₃C-R; ~0.9-3.3 ppm), methylene (H₂CH₂-RR‘; ~1.3-2 ppm), and methine (HC-R,R’,R”; ~1.4-4.5 ppm).

Figure 2.11 High resolution mass spectrum (positive ion) from SPE-DOM collected at 250 m, Station Aloha, near Hawaii. The mass spectrum of the infused sample (bottom) shows a complex distribution of ions with reoccurring mass differences of Δm = 14.0156 due to methylene homologues (Δ CH₂). Expansion of the series centered at 377 Da (upper left) shows a second homologous series with Δm = 2.0157 (H₂). Further expansion (upper right) shows the high mass resolving power that can be achieved, which allows for the assignment of elemental formulae, and the distinction of formulae of the same nominal mass. Here ions that differ by the substitution of CH₄ (16.0313 Da) for O (15.9949) yield a mass difference of 36.4 mDa.

Figure 2.12 van Krevelen plot of data in Figure 2.11 from 250 m, Station ALOHA. Each dot represents an ion for which a unique molecular formula could be assigned from the exact mass. Here the elemental ratios H/C and O/C calculated from molecular formulae are plotted and compared with elemental H/C and O/C ratios from classes of known biochemical (lipids, proteins, and carbohydrates). Most SPE-DOM formulae plot outside the region of known biochemicals, perhaps due to extensive degradation and transformation of organic matter during DOM formation, or from unknown DOM precursors. A distinct group of ions with H/C < 1 and O/C< 0.2 appear in some SPE-
DOM spectra; these have been assigned to polycyclic aromatic compounds (PCAs) of thermogenic origin (see text).

Figure 2.13 HPLC analysis of benzene polycarboxylic acids (BPCAs) derived from SPE-DOM. Treatment of SPE-DOM with nitric acid at high temperature degrades polycyclic aromatic compounds (top) into a series of benzoic acids that can be measured at high sensitivity by HPLC (bottom) and HPLC-MS.

Figure 2.14 The microbial loop of labile DOM cycling. Carbon dioxide is fixed by photoautotrophs (yellow and green cells, left), which excrete about 10% of total fixed carbon as DOC. Grazing (bottom left, right) and cell lysis after viral infection (upper left, right) release additional labile DOM to be consumed by heterotrophic bacteria and archaea (blue cells, right). The microbial loop consumes labile DOM, converting it to carbon dioxide and nutrients. Heterotrophs also release essential vitamins, siderophores, and other organic compounds that facilitate the growth of photoautotrophs. Figure courtesy of Dr. Jamie Becker.

Figure 2.15 $^{13}$C NMR of DOM recovered by RO/ED (black trace) superimposed on spectra of HMWDOM (ultrafiltration; grey trace) and SPE-DOM (red trace) from surface water. The spectra of HMWDOM and SPE-DOM are scaled to the RO/ED peak at 70 ppm. To a good approximation, the distribution of carbon functional groups in the RO/ED sample is a mixture of acylated polysaccharides (APS), which dominate the $^{13}$C NMR spectra of HMWDOM, and hydrophobic humic substances/CRAM that dominate the $^{13}$C NMR spectra of SPE-DOM. Figure redrawn from Koprivnjak et al., 2009.
Figure 1
Figure 2
Figure 4
Figure 5

Arabinose

Fucose

Mannose

N-acetyl glucosamine

3-Methyl rhamnose

3-deoxyglucosamine
Figure 8

Amino Acid

Relative Mole %

- THAA (dissolved)
- Plankton
- Suspended Particles
- Sinking Particles
Figure 11

\[ \Delta m = 2.0157 \text{ Da} \]

\[ \Delta m = 36.4 \text{ m Da} \]

\[ \Delta m = 14.0156 \text{ Da} \]

\[ C_{17}H_{24}O_{8}\text{Na} \]
\[ 379.1369 \text{ Da} \]

\[ C_{16}H_{20}O_{9}\text{Na} \]
\[ 379.1005 \text{ Da} \]

\[ C_{18}H_{28}O_{7}\text{Na} \]
\[ 379.1732 \text{ Da} \]
Station Aloha (250)

Figure 12
Figure 13

Pyrene $\xrightarrow{\text{HNO}_3, \text{heat}}$ Hemimellitic acid $\xrightarrow{}$ Prehnitic acid

![Diagram showing the chemical structures of pyrene, hemimellitic acid, and prehnitic acid](image-url)
Figure 14
2.6 References


Lindroth P. and Mopper K. (1979) High performance liquid chromatographic
determination of subpicomole amounts of amino acids by precolumn fluorescence


*Science*** **280**, 1911–1913.

Achterberg E. P. (2011) Production of siderophore type chelates in Atlantic Ocean

dissociation of three groups of hydroxamate siderophores: ferrioxamines,

McCarren J., Becker J. W., Repeta D. J., Shi Y., Young C. R., Malmstrom R. R.,
reveal microbes and metabolic pathways associated with dissolved organic matter

McCarthy M. D., Beaupré S. R., Walker B. D., Voparil I., Guilderson T. P. and Druffel


McCarthy M., Hedges J. and Benner R. (1996) Major biochemical composition of


