Deciphering Ocean Carbon in a Changing World

Mary Ann Moran*[,1], Elizabeth B. Kujawinski*[,2], Aron Stubbins*[,3], Rob Fatland*[,4], Lihini I. Aluwihare[5], Alison Buchan[6], Byron C. Crump[7], Pieter C. Dorrestein[8], Sonya T. Dyhrman[9], Nancy J. Hess[10], Bill Howe[11], Krista Longnecker[2], Patricia M. Medeiros[1], Jutta Niggemann[12], Ingrid Obernosterer[13], Daniel J. Repeta[2], Jacob R. Waldbauer[14]

Affiliations:

[1] – Department of Marine Sciences, University of Georgia, Athens, GA 30602, USA
[2] – Department of Marine Chemistry & Geochemistry, Woods Hole Oceanographic Institution, Woods Hole MA 02543, USA
[3] – Skidaway Institute of Oceanography, Department of Marine Sciences, University of Georgia, Savannah, GA 31411, USA
[4] – Microsoft Research, Redmond, WA 98052, USA
[5] – Geosciences Research Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093 USA
[6] – Department of Microbiology, University of Tennessee, Knoxville, TN 37996, USA
[7] – College of Earth, Ocean, and Atmospheric Science, Oregon State University, Corvallis, OR 97331, USA
[8] – Departments of Pharmacology, Chemistry, and Biochemistry, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92037, USA
[9] – Department of Earth and Environmental Science and the Lamont-Doherty Earth Observatory, Columbia University, Palisades NY 10964, USA
[10] – Environmental Molecular Science Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352, USA
[11] – Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195, USA
[12] – Research Group for Marine Geochemistry (ICBM-MPI Bridging Group), Institute for Chemistry and Biology of the Marine Environment (ICBM), University of Oldenburg, Germany
[13] - CNRS, Sorbonne Universités, UPMC Univ Paris 06, Laboratoire d'Océanographie Microbienne (LOMIC), Observatoire Océanologique, F-66650, Banyuls/mer, France
[14] - Department of the Geophysical Sciences, University of Chicago, Chicago, IL 60637, USA

*Co-corresponding authors
MAM: mmoran@uga.edu
EBK: ekujawinski@whoi.edu
AS: aron.stubbins@skio.uga.edu
RF: rob5@uw.edu

For submission to: PNAS (Perspective); Classification: Biological Sciences

Abstract

Dissolved organic matter (DOM) in the oceans is one of the largest pools of reduced carbon on Earth, comparable in size to the atmospheric CO_2 reservoir. A vast number of compounds are present in DOM and they play important roles in all major element cycles, contribute to the storage of atmospheric CO_2 in the ocean, support marine ecosystems, and facilitate interactions between organisms. At the heart of the DOM cycle lie molecular-level relationships between the individual compounds in DOM and the members of the ocean microbiome that produce and consume them. In the past, these connections have eluded clear definition because of the sheer numerical complexity of both DOM molecules and microorganisms. Emerging tools in analytical chemistry, microbiology and informatics are breaking down the barriers to a fuller appreciation of these connections. Here we highlight questions being addressed using recent methodological and technological developments in those fields and consider how these advances are transforming our understanding of some of the most important reactions of the marine carbon cycle.

Significance statement

Marine dissolved organic matter is the pool through which one quarter of all photosynthesis on Earth is cycled. This large and biologically-active organic matter reservoir lies at the base of marine food webs and at the heart of the ocean carbon cycle. Marine DOM is made up of tens to hundreds of thousands of distinct molecules that are shaped by an equally diverse marine microbiome comprised of bacteria, archaea, microeukaryotes, and viruses that produce and consume its components. Most links between microbes and DOM have defied characterization thus far due to the inherent complexity of each, but recent advances in chemistry, biology and informatics are shedding light on microbe-DOM relationships and, in turn, promoting new understanding of the global carbon cycle.

\body

Introduction

The global cycling of carbon supports life on Earth and affects the state of the biosphere within which humans reside. Industrial processes are now altering the balance of this natural cycle by adding fossil carbon to the contemporary atmosphere and changing our climate (1). Marine dissolved organic matter (DOM) is central to the current and future global cycle, storing as much carbon as the current atmospheric CO_2 reservoir (2) (Fig. 1).

Flux of carbon through the marine DOM pool is mediated largely by microbial activity. Yet the intertwined relationships between the molecules making up the DOM pool and the ocean microbes that process them remain poorly characterized. In short, the complexity of each has defied easy characterization, and fundamental interactions have been necessarily over-simplified to yield a scientifically tractable framework. The principles of organization and interactions between ocean microbial communities and DOM have parallels in other complex ecosystems such as mammalian microbiomes, soils, rhizospheres, extreme environments, and the built environment. Thus progress in mapping microbe-DOM interactions in the oceans will enhance knowledge across seemingly disparate fields, culminating in a better understanding of element cycling in Earth's varied ecosystems.

Recent advances in chemistry, microbiology and data science have directly addressed the complexity of DOM cycling in marine environments and led to a reexamination of basic concepts. A revolution in DNA sequencing technology (3), advances in mass spectrometry (4-6), new approaches to identify metabolites from genome sequences (7), the growth of informatics (8, 9), and the building of knowledge and analysis cyberinfrastructures (10-12) are key tools already in place or in development. As a result, the DOM pool is now known to conservatively consist of tens to hundreds of thousands of different organic molecules (13), for which formulas are rapidly emerging (14). Meanwhile, the ocean microbiome has been estimated to consist of more than a hundred thousand different bacterial, archaeal, and eukaryotic taxa (15, 16) with diverse ecological and metabolic strategies for producing and consuming fixed carbon (17-19). Until recently, major gains in understanding ocean carbon cycling have moved largely along independent lines within the fields of biology and chemistry. Now, it is at the confluence of these disciplines, enabled through innovative data science, that transformative advances are being made (Fig. 2).

Here we present six fundamental questions in marine biogeochemistry that are benefiting from integrated research strategies. The questions are organized along a general gradient in apparent DOM reactivity that is based on persistence under typical ocean conditions (2). "Labile"

DOM refers to the molecules that are consumed by microbes within hours or days of production (Fig. 1). “Semi-labile” DOM is less reactive and persists in the surface ocean for weeks to years. "Refractory" DOM is the least biologically reactive and circulates through the major ocean basins on millennial time scales. Although all three operational categories occur throughout the ocean, their relative importance loosely corresponds to a depth gradient. In the surface ocean, the photosynthesis of organic molecules from CO_2 by phytoplankton (i.e., primary production) is the source of most of the ocean's labile and semi-labile DOM. Semi-labile DOM persists long enough to be transported to moderate ocean depths (100s of meters below the surface) before it is metabolized (20). Refractory DOM has its strongest signature in the deep ocean (depths greater than 1000 meters) (2, 21). The linkages among individual molecules and microbes that culminate in the global carbon cycle and give rise to the DOM reactivity spectrum lie at the foundation of the questions posed here.

Which compounds represent the largest conduits of carbon flux through the labile marine DOM pool?

Each year in the surface ocean, ~20 Gt of carbon recently fixed into organic matter by phytoplankton photosynthesis is rapidly taken up by heterotrophic bacteria (1 Gt = one giga tonne or 1 x 10^{15} grams). For perspective, the current annual increase in the atmospheric CO_2 pool is 4 Gt C (22) and annual processing of refractory marine DOM is <0.2 Gt C. Many of the labile compounds mediating this brisk and quantitatively important carbon flux into the microbial food web are thought to have half-lives on the order of minutes and concentrations in the picomolar range (23). The very characteristics that define highly labile compounds make their study extremely challenging.

Because phytoplankton cells are rich in proteins and carbohydrates, and these polymers are typically degraded extracellularly into oligomers or monomers prior to transport into bacterial cells, early research on biologically labile DOM focused primarily on rates and kinetics of amino acid and sugar uptake (24-26). Today, a wealth of new data coming largely but not exclusively from the 'omics tools (genomics, transcriptomics, proteomics, and metabolomics) suggests that a much wider variety of molecules participate in the rapid heterotrophic DOM flux. For example, gene expression studies in both the ocean and laboratory indicate that labile DOM can take the form of mono- and dicarboxylic acids (27, 28), glycerols and fatty acids (27, 29), and the nitrogen-containing metabolites taurine, choline, sarcosine, polyamines, methylamines, and ectoine (27, 30, 31). One-carbon compounds such as methanol (27, 29, 31) as well as several sulfonates (32) have

recently been added to the list. Chemical analyses concur; photosynthate released directly from phytoplankton is highly complex, consisting of hundreds of different compounds (33, 34). This "dissolved primary production" — the material released from living phytoplankton — supports a major fraction of labile carbon flux in the surface ocean (35).

Complexity in the composition and concentration of labile DOM presents an ecological opportunity for microbes but an analytical challenge for chemists. Substrates used by heterotrophic bacteria will not accumulate if their demand is higher than their supply; therefore, the most important biologically labile molecules are inherently difficult to recognize against the chemical background of organic compounds in seawater. For example, monomeric amino acids and sugars have concentrations below one billionth of a gram per liter, which is at or below the limits of quantification in marine waters (36, 37). However, detecting low-concentration high-flux compounds has recently become more tractable with methodological advances in chemistry (e.g., better separation methodologies, sensitivity, accuracy, and resolving power; 5, 38, 39), biology (e.g., deducing key substrates from transcriptome analysis; 28, 29, 40), and cyberinfrastructure (e.g., determining patterns of DOM-bacterial interaction networks; 41). The complementarity of these research fields is key to identifying this massive yet all but invisible flux in the ocean's active carbon cycle (Fig. 2).

How are element cycles linked through marine DOM?

In addition to driving major fluxes of carbon, microbial production and consumption of DOM in the surface ocean also plays a central role in the cycling of nitrogen (N), phosphorus (P) and sulfur (S), along with micronutrients such as iron, cobalt, nickel, and zinc. Molecules within the marine DOM pool that contain N, P or S include amino acids and proteins (42), nucleotides and nucleic acids (43), various osmolytes (44), siderophores (45), vitamins (46), and primary metabolites.

Advances in understanding the fate of these diverse components of DOM have occurred in spite of the fact that extraction of element-specific compound classes quantitatively from seawater remains a challenge, and that characterization of the myriad biological systems that support uptake and transformation of N-, P-, and S-containing organic compounds is daunting. Two main lines of scientific inquiry have motivated progress. The first is a long-standing question in oceanography on the role of organic forms of N, P and S in alleviating nutrient stress for ocean microbes. This question is particularly relevant in oligotrophic oceans where inorganic nutrients (that is, nitrate, nitrite, ammonium, phosphate, and sulfate) are perennially limiting or energetically expensive to reduce to biologically active forms. Microbes that use organic nutrients may have a significant

advantage over those that cannot. With 'omics data as an influential driver, substantial progress is being made in understanding the microbial production and use of organic nutrients. Phosphorylated organic compounds, for example, provide a source of inorganic phosphate after cleavage by phosphatases and nucleotidases (47, 48). Phosphonates are synthesized by marine cyanobacteria, archaea, and other microbes (49-52) and subsequently consumed both by microbial autotrophs (53, 54) and heterotrophs (55). Nitrogen stress is lessened in the oligotrophic ocean by use of urea and cyanate (56, 57). The fate of these dissolved organic N, P, and S molecules in seawater represents a confluence of Earth's element cycles.

The second issue stimulating research into organic N, P, and S-containing compounds is their use as biochemical intermediates and co-factors by auxotrophic microbes (i.e., those unable to synthesize metabolites critical for their own growth). Auxotrophy in the ocean is thought to reflect a microorganism's evolutionary positioning along the trade-off between the expense of biosynthesis of complex molecules on the one hand, and the risk of relying on neighbors on the other (58). For instance, bacteria in the Pelagibacterales are missing the genetic capability for using extremely abundant sulfate, and instead scavenge organic sulfur from seawater (59, 60); this is truly remarkable for what is arguably the most successful heterotrophic microbial group in the ocean. Many marine phytoplankton with critical roles in global carbon fixation have lost biosynthetic pathways for N-, P- and S-rich vitamins such as B_1 and B_{12} (61) and must scavenge them from the DOM pool. Genomic data have been ideal for learning which marine microbes depend on the DOM pool for energetically expensive biomolecules, while metabolomics advances have detected dilute components of DOM that were previously not measurable (39, 46).

How do microbe-microbe relationships influence DOM?

Early studies of the marine microbial food web revealed a major role for trophic interactions in the formation and flux of DOM. A surprising 20-50% of microbial biomass is turned over each day in the ocean by viral infection (62), releasing intracellular organic matter into surrounding seawater (63-65). Similarly, protistan grazing on bacteria and phytoplankton converts up to 30% of ingested carbon to dissolved form (66). Protists also directly consume DOM (67) in this intricate network of microbial predation.

Our knowledge of the DOM molecules that arise from or facilitate interactions between marine microbes is growing. Metabolomics approaches have revealed that viral infection increases the concentrations of N-rich metabolites in infected bacteria (38). In this cycle-within-a-cycle, a portion of the organic matter initially assimilated from the DOM pool into bacterial biomass is

returned to the DOM pool as viral lysate (62) but is enhanced in N relative to metabolites of non-infected cells (38). New categories of non-predatory microbial alliances that release organic compounds into seawater are also being recognized. These include molecules in microbial cytosols and exudates (39, 68) that serve as substrates, signaling molecules, and allelochemicals to neighboring microbes (69, 70). Genes able to mediate microbial interactions have also been uncovered (71, 72), including a high prevalence of virulence gene homologs in marine bacteria and archaea that could facilitate direct contact with eukaryotic plankton (73). The conditions under which genes mediating microbial interactions in the ocean are expressed are now better understood because of metatranscriptomic surveys (30, 32). Examples include the alteration of marine phytoplankton growth rates by bacterial release of phenylacetic acid (74) and indole-acetic acid (72), and the modulation of bacterial quorum sensing molecules (75) and antibiotic production (76) by phytoplankton. Global patterns of marine plankton co-occurrences can be better explained by factors involving microbial interactions (such as grazing, viral infection, and parasitic relationships) than by environmental conditions (77). Thus it has become clear that compounds released into the DOM pool by ocean microbes are considerably more chemically diverse than predicted from the composition of plankton biomass, at least in part because many are synthesized for roles occurring beyond the cell wall.

The full consequences of microbe-microbe interactions depend to a large extent on factors such as cell encounter frequencies in seawater and life history traits of the microbial participants (78). Thus their prediction must also rely on modeling approaches that consider small- and large-scale dynamics and feedbacks in ocean waters. As examples, a heterotrophic bacterium will experience higher DOM concentrations when associated with a particle compared to when it is free-living (78, 79) and viral-host interactions can at the same time kill individual cells while stimulating overall ecosystem productivity (80). New generations of biogeochemical models are explicitly incorporating 'omics-derived data (81, 82) to more directly link microbes and the fate of marine DOM.

How many metabolic pathways are required for the bacterial transformation of marine DOM? While there is currently no way to know the full biochemical diversity behind microbial processing of DOM, headway is being made. Early genomic studies addressing this question typically focused on transporter systems because they mediate the first essential step in utilization of DOM by heterotrophic bacteria. Transporters involved in organic compound uptake have been reported to account for 13% of expressed genes (28) and 35% of expressed peptides (83) in marine

bacterial communities. Over 100 protein families predicted to function in the uptake of organic compounds from seawater have been described in metagenomic, metatranscriptomic and metaproteomic datasets (28, 84, 85).

Yet assigning an exact function to microbial genes remains a stubborn hurdle to more effective use of genomic databases to address DOM transformations. In the case of transporter genes, many are poorly annotated with regard to substrate specificity, and consequently are assigned only to broad categories (such as "branched chain amino acid" or "carboxylate" transporter) based on homology to a limited number of experimentally characterized genes. Making matters worse, transporters classified into the same protein family may mediate uptake of different substrates (86); a single transporter can have multiple substrates (87); and higher molecular weight DOM is assimilated through generic, and therefore uninformative, transporter systems following hydrolysis at the cell surface (88). In the case of catabolic genes, those encoding conserved central metabolic pathways are generally well characterized, but pathways for upstream reactions that ultimately feed into central metabolism, the linchpins of many essential biogeochemical transformations, are poorly known. When we are ultimately successful in identifying a new biogeochemically-relevant gene, it is often the case that it was initially annotated with a misleading or uninformative function (89-91).

Substantive improvement in characterization of microbial genes is widely acknowledged as a central goal in biology. One approach to annotation of genes relevant to ocean carbon cycling is the use of model microbial systems for which transcription patterns, metabolomics, protein expression, and genetic systems can be leveraged. For example, the first marine bacterial gene mediating catabolism of dimethylsulfoniopropionate (DMSP), a compound known for 50 years to be critical in the global sulfur cycle, was found in 2006 by generating transposon mutants of a marine bacterium (92) and the first phytoplankton gene mediating synthesis of DMSP was found in 2015 by shotgun proteomics of phytoplankton isolates (93). A full suite of DMSP gene discoveries is now enabling studies of the dominant transformation pathways and their regulation in the ocean (94, 95). Similarly, the genes mediating marine bacterial transport and metabolism of organic compounds such as sulfonates (90, 96), ectoine and hydroxyectoine (97), and methylamines and choline (31, 98, 99) have been recently elucidated through model organism systems.

This type of characterization work is slow due both to the small fraction of marine bacteria amenable to culturing and the challenges of developing genetic systems for them. A culture-independent twist involves cloning DNA from marine environments into laboratory strains and performing screens to identify genes conferring a function of interest. Examples include an early

effort that identified chitin degradation genes in marine microbial community DNA (100) and a recent study that discovered genes for use of a novel phosphonate (55).

The expansion of cyberinfrastructure capabilities has opened up possibilities of using pattern mining of combinatorial datasets (DOM and metabolite composition, microbial genes and transcript inventories) to generate hypotheses regarding gene function. This approach is already having success in secondary metabolite research (7, 101). Such efforts will be particularly informative when guided by knowledge of which unknown protein families are ubiquitous in genomes of ocean microbes, or demonstrate phylogenetic coherence, or show biogeographic patterns. Answering the question of how many metabolic pathways are required for the bacterial transformation of marine DOM is perhaps an impossible task, but identification of a subset of pathways that mediate important fluxes of dissolved compounds through the oceanic carbon reservoir is steadily pushing understanding forward.

Why does semi-labile DOM accumulate in the surface ocean?

Semi-labile DOM is operationally defined as the dissolved organic compounds that accumulate in surface waters over time frames of weeks to years but then disappear once exported to depth (20). Why, exactly, these molecules resist degradation in the surface ocean where heterotrophic microbes are often limited for substrates and nutrients remains a mystery. As a substantial and temporally-stable component of DOM, the semi-labile pool affects the overall rate of carbon turnover in the oceans (Fig. 1). Thus illuminating its composition and identifying the metabolic pathways that can degrade it are important for predictive understanding of carbon sequestration (that is, the transfer of excess carbon from the atmosphere into long-term storage in the ocean).

New data are beginning to untangle the factors that co-vary with semi-labile DOM and depth in the ocean, and thereby helping to understanding its fate. Microbial diversity is lower in surface than in deep waters, which suggests that a limited genetic repertoire in surface heterotrophs might restrict degradation of certain compounds (15). At the species level, oligotrophic bacteria such as Pelagibacterales dominate open ocean gyres where semi-labile DOM accumulates, and these cells typically have small genomes with fewer and less varied transporters and catabolic pathways (102). Recent experiments with the marine bacterium *Alteromonas*, harboring a substantially larger genome than the Pelagibacterales, showed that while this one strain can degrade the labile fraction of marine DOM in a period of days, an amount of DOM equivalent to the semi-labile pool remained untouched. Instead, the full microbial community was needed to degrade the semi-labile DOM (103). Earlier studies had shown that the addition of both

labile DOM and inorganic nutrients is needed to degrade semi-labile DOM (104), signifying a complex relationship between DOM accumulation, microbial diversity, and the availability of nutrients and co-metabolites.

The chemical and optical signatures of seawater also differ from the deep ocean background in locations where semi-labile DOM accumulates, indicating that this material is compositionally distinct from labile or refractory DOM. For instance, fluorescence signals indicative of dissolved proteins are elevated and signatures of carbohydrates and aliphatic material are enriched in semi-labile DOM relative to the signatures of deep ocean DOM (105, 106). Nevertheless, extraction protocols are insufficient and chemical understanding too limited to physically isolate semi-labile DOM from seawater at this time. Instead, indirect experiments such as time-series studies (107), long-term incubations (103), and the isolation of representative microbes (108) are being used to address first-order questions regarding the molecular composition of this enigmatic DOM pool and the metabolic pathways by which it is degraded. New data analysis methods are also helping to parse small but crucial biological signals from these complex data sets.

How refractory is deep-ocean DOM and why does it persist?

The deep ocean represents a challenging ecosystem to study because of its remoteness, the low concentrations of organic molecules, and the slow rates of microbial metabolism at high pressure and low temperature. Yet this is the repository for over 70% of the carbon sequestered in DOM and a major reservoir in the global carbon cycle (Fig. 1). Bulk radiocarbon dating indicates that deep ocean DOM has an average age of 6,000 years (109). More recent radiocarbon techniques showed that the apparent ages of individual molecules are not normally distributed around this average. Instead, different reactivity pools were identified representing both semi-labile (radiocarbon enriched) and refractory (radiocarbon deplete) pools within the DOM, with the most deplete fraction having a radiocarbon age of ~12,000 years (21).

That these energy-rich molecules exist for millennia in the deep ocean is a paradox that seems to contradict the laws of thermodynamics – why in a marine environment rich in other necessities for life would microbes fail to use such a large reservoir of organic carbon? One line of reasoning posits that this pool contains inherently biologically recalcitrant molecules. For example, condensed polycyclic aromatics generated by processes such as wildfire and biomass burning on land accumulate throughout the deep ocean (110, 111) and have radiocarbon ages exceeding those of other DOM pools (112). These molecules are susceptible to photodegradation because of their aromatic functional groups, suggesting that upwelling of deep waters to the surface during ocean

circulation may determine the half-life of photochemically active yet biologically refractory compounds (113-115). The majority of deep ocean refractory DOM, however, likely represents the accumulation of metabolic products of ocean microbes that are refractory to further biological degradation (116), a phenomenon termed the "microbial carbon pump" (117). It is not clear if the microbial carbon pump generates inherently refractory molecules from labile forms, or if labile DOM is diversified by the pump until each molecule is present at vanishingly low concentrations. In the former case, refractory DOM would consist of a pool of survivor molecules that are biologically intractable and enriched at depth. In the latter case, refractory DOM would represent a highly diverse suite of compounds each at its limiting concentration of metabolic utility (118-121).

Conducting the laboratory and field experiments to test current theories of the nature of refractory deep-ocean DOM is proving to be both enlightening and challenging. Incubation experiments seeking to measure changes in DOM concentrations and chemistry under conditions that mimic the deep ocean are hampered by inherently low rates of net carbon turnover and analytical techniques that provide limited structural resolution of resistant molecules (121-123). Further, refractory organic matter is defined based on its lifetime in the ocean (2) rather than on inherent chemical structures, making for an elusive experimental target. Extraction of DOM from seawater, a prerequisite to most analytical methods, does not presently yield all the compounds dissolved in seawater (124). Thus, our view of the molecular composition of DOM remains restricted to the fraction that can be physically isolated and analyzed. On the biological side, the percentage of microbial genes with no known function increases with depth in the ocean (15). The metabolic pathways that degrade refractory molecules may be hidden within these unannotated genes with no analogs in known metabolic pathways. When the question of why deep ocean DOM persists is finally resolved, the answer is likely to be a combination of concentration, chemical structure, bioenergetics, and microbial diversity.

A final notable aspect of the deep ocean ecosystem is that while it is home to the largest reservoir of refractory DOM, it also harbors labile and semi-labile molecules. Multiple lines of evidence have recently revealed labile DOM-microbe interactions far below the photic zone (21, 125-128). For example, carbon isotopic analysis of microbial DNA confirms the incorporation of modern organic matter into microbial biomass in the ocean depths (129). Release from sinking particles is the primary recognized source of modern DOM at depth. Indeed, the dissolved organic compounds liberated by microbes from sinking particles are now thought to fuel up to 90% of carbon cycling in the deep ocean (125, 129-133).

The next step: Prototypical molecules of the marine carbon cycle

An opportunity to identify a broader range of molecular currencies of the marine carbon cycle can be found at the intersection of marine chemistry and 'omics methodologies, in the context of developments in informatics. Admittedly, successful identification of even several hundred new molecules seems a trivial advance stacked against the enormous chemical diversity of seawater organic matter. Yet microbial ecologists would nearly unanimously agree that genome sequences of just 175 marine bacteria (134) out of the hundred thousand taxa present in seawater fueled a revolution in our understanding of element cycling in the ocean. Indeed, many of the recent DOM advances discussed here were directly enabled by foundational data on the genomes of marine plankton (28, 29, 31-33, 38, 39, 70). A corresponding suite of model organic compounds representative of those cycling through the world's oceans will provide more muscle for the job of unraveling pathways of carbon flux.

Two categories of prototypical molecules are of particular interest in this endeavor. The first is molecules rapidly produced and metabolized by marine microbes in the fast loop of labile DOM, including biogeochemical intermediates and signaling compounds relevant to organic matter flux. The second is molecules from less labile components of marine DOM that will improve understanding of why molecules are biologically refractory and what characteristics determine their half-life in the ocean. Identification of prototypical compounds will then lead to methodologies for their analysis in bulk seawater and microbial metabolomes, and to synthesis and labeling for flux studies.

Already, modern targeted chemical workflows are enabling quantification of intermediates of biogeochemical cycling. Correspondingly, non-targeted workflows are helping us to discover new molecules we didn't know to look for (33, 39, 70, 105, 135). Newly developed informatics approaches are supporting data mining across multiple studies and systems (136, 137) (see Box 1). 'Omics data are allowing us to use microbes as biosensors for the compounds being synthesized, assimilated, and metabolized in the ocean microbiome (27, 28, 138-140). Genetic systems are assigning substrates to uncharacterized genes through knockouts and heterologous expression (31, 76, 92). All of these tools and others on the horizon will expand our knowledge of the organic compounds produced and transformed by microbes of the ocean.

Conclusions

Exciting discoveries have been moving the needle on our understanding of the marine microbe-DOM network over the past decade. Successes include improved knowledge of the organic

compounds through which nearly a quarter of net global photosynthesis passes within days of fixation, a grasp of the chemical formulas of compounds that persist for tens of thousands of years in the ocean, knowledge of how organic forms of limiting nutrients take part in element cycles, and realization of the crucial roles of marine microbial interactions in Earth's biogeochemical cycles. More discoveries are in the pipeline, helped by innovation in high-throughput methodologies and effective cyberinfrastructures. The next decade will continue this period of rapid learning, both in ways that we glimpse already (through growing accessibility of metabolomics, the speed and lowered cost of next-generation sequencing, the development of efficient screening tools for gene function) and from directions not yet predictable.

The DOM-microbe complexity challenge has synergies with other areas of science where the chemical foundations of microbial community function are crucial. Microbiome studies, for example, have the same scientific goals of discovering, identifying, and quantifying molecules that link a genome-encoded potential with a realized metabolic and ecological function. Annotation of gene function likewise cuts across many fields and organisms. A compelling example is the recent discovery of the genetic basis of bacterial degradation of the sulfolipid component of photosynthetic membranes, based on studies conducted with bacteria from soil (90), coastal seawater (32), and the human gut (89). Another example is the development of cross-discipline databases for metabolite annotation, including the use of crowd-sourcing to solve common problems in compound identification (gnps.ucsd.edu). The new classes of data and types of methodologies being developed to explore both molecules and microbes will be necessary to predict carbon cycle response to challenges ranging from oil spills to climate change (see Box 2).

Box 1: Cyberinfrastructure

The merging of chemical and microbiological data for resolving microbe-DOM interactions in the ocean is being enabled by advances in data management capabilities and systems, collectively referred to as cyberinfrastructure. In genomics research, the core cyberinfrastructure method is well established: sequence databases are searched for homology using tools such as BLAST, and then analyzed for taxonomy and function. Analogous datasets and cyberinfrastructure are now emerging that can be applied to investigate DOM chemistry, for example MetaboLights (http://www.ebi.ac.uk/metabolights/) and GNPS (http://gnps.ucsd.edu) which emphasize mass spectrometry knowledge capture and dissemination using social networking. In concert with data accessibility (141), new or existing infrastructure can be specifically dedicated to the growing needs of the DOM community.

Well-engineered data systems adopted by collaborating scientists are the key to keeping up with the burgeoning capacity to generate data. For microbe-DOM research, such systems will require coordinated cyberinfrastructure elements that include descriptions of chemical composition; inventories and interpretation of transcripts, genes, and proteins; and data that are curated and searchable. Publication of open-access datasets must be easy and rewarded. Open source tools for data reduction and inter-comparison, such as PARAFAC for fluorescence data (142), must be developed. By this approach, processing tasks once considered difficult can be automated. Validation, provisions for searchable metadata, provenance, repeatability, and archiving are all considerations that weave into robust cyberinfrastructure development.

Two related elements of cyberinfrastructure design are managing data volume and making data more available. This latter refers both to researchers not directly involved in acquisition and to questions that are not yet anticipated. For example, a field scientist could generate synoptic assays of the near-surface microbiome – DOM systems at a rate of one snapshot every few minutes over a period of weeks, actively tracking the biogeochemical pathways of the ocean. Well-designed cyberinfrastructure would make the resulting data discoverable, explorable, and queryable by other scientists, in addition to performing data reduction and organizational tasks at the many-Terabyte scale. As we envision scientists being rewarded for proliferating public data and software, so too should cyberinfrastructure developers be rewarded for building data systems that reduce analysis times from months to minutes and for coordinating with data discovery mechanisms in the scientific community.

Box 2: Microbe-DOM climate responses

Climate change effects on ocean systems are being manifested as global shifts in temperature, seawater pH, sea level, circulation patterns, oxygen content, and nutrient and DOM loading from land. Marine ecosystems are also being affected regionally by coastal eutrophication, invasive species, and habitat degradation. Today, oligotrophic subtropical gyres are regions of DOM accumulation (143) and export (144). The predicted growth in the areal extent of gyres in the future, evidenced by a 56% increase of the North Atlantic gyre wintertime area between 1998 and 2006 (145, 146), may therefore increase net oceanic DOM production. Yet experimental studies suggest rising temperatures and ocean acidification will increase bacterial DOM consumption (147, 148), while the same drivers may reduce formation of colloids and microgels from DOM (149). Thus, whether or not the future ocean will experience greater accumulation of DOM or alterations in its chemical composition (150) is still unclear.

Climate change is also predicted to alter the distribution and composition of marine phytoplankton communities and create new physical regimes that shift longstanding chemical distributions, throwing together microbes and carbon forms with limited evolutionary history. The emergence of new high temperature oceanic biomes, currently rare regions where mean sea surface temperatures exceed 31°C, is projected to establish more than 25 million km^2 of altered ocean by 2100 (145). Whether microbes inhabiting these and other new niches will interact with DOM as analogs of current assemblages is unknown. Discovery and prediction of microbe-DOM linkages as they react to and shape the future ocean will rely heavily on the tools and concepts discussed in this perspective.

Acknowledgements

Formative discussions for this perspective occurred at a workshop entitled "Linking Marine Microbes and the Molecules of Dissolved Organic Matter", held in New York City in November 2014. The workshop was supported by the Gordon and Betty Moore Foundation and Microsoft Research Corporation. Additional support was provided by National Science Foundation grants OCE1356010, OCE1154320, and OCE1356890, and Gordon and Betty Moore Foundation Grant #3304. We thank Jack Cook for graphics expertise.

References

1. Doney SC, *et al.* (2012) Climate change impacts on marine ecosystems. *Ann Review Mar Sci* 4:11-37.
2. Hansell DA (2013) Recalcitrant dissolved organic carbon fractions. *Ann Rev Mar Sci* 5:421-445.
3. DeLong EF & Karl DM (2005) Genomic perspectives in microbial oceanography. *Nature* 437(7057):336-342.
4. Koch BP, Ludwichowski KU, Kattner G, Dittmar T, & Witt M (2008) Advanced characterization of marine dissolved organic matter by combining reversed-phase liquid chromatography and FT-ICR-MS. *Mar. Chem.* 111(3-4):233-241.
5. Bouslimani A, Sanchez LM, Garg N, & Dorrestein PC (2014) Mass spectrometry of natural products: current, emerging and future technologies. *Nat. Prod. Rep.* 31(6):718-729.
6. Kujawinski EB, Hatcher PG, & Freitas MA (2002) High-resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids: improvements and comparisons. *Anal. Chem.* 74(2):413-419.
7. Cimermancic P, *et al.* (2014) Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* 158(2):412-421.
8. Dührkop K, Shen H, Meusel M, Rousu J, & Böcker S (2015) Searching molecular structure databases with tandem mass spectra using CSI: FingerID. *Proc. Natl. Acad. Sci. USA.* doi:10.1073/pnas.1509788112.
9. Watrous J, *et al.* (2012) Mass spectral molecular networking of living microbial colonies. *Proc. Natl. Acad. Sci. USA* 109(26):E1743-E1752.
10. Amaral-Zettler L, *et al.* (2010) A global census of marine microbes. *Life in the world's oceans: diversity, distribution and abundance*, ed McIntyre AD (Wiley-Blackwell, Oxford, UK), pp 223-245.
11. Haug K, *et al.* (2012) MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Res.* 41(Database issue):D781–D786.
12. Sun S, *et al.* (2010) Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis: the CAMERA resource. *Nucleic Acids Res.* 39(Database issue):D546-D551.
13. Kim S, Kramer RW, & Hatcher PG (2003) Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. *Anal. Chem.* 75(20):5336-5344.
14. Hertkorn N, *et al.* (2006) Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Acta* 70(12):2990-3010.
15. Sunagawa S, *et al.* (2015) Structure and function of the global ocean microbiome. *Science* 348(6237):1261359.
16. Gibbons SM, *et al.* (2013) Evidence for a persistent microbial seed bank throughout the global ocean. *Proc. Natl. Acad. Sci. USA* 110(12):4651-4655.
17. Ganesh S, Parris DJ, DeLong EF, & Stewart FJ (2014) Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J* 8(1):187-211.
18. Marshall KT & Morris RM (2013) Isolation of an aerobic sulfur oxidizer from the SUP05/Arctic96BD-19 clade. *ISME J* 7(2):452-455.
19. Moran MA, *et al.* (2004) Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* 432(7019):910-913.
20. Hansell DA & Carlson CA (1998) Net community production of dissolved organic carbon. *Global Biogeochem. Cycles* 12(3):443-453.
21. Follett CL, Repeta DJ, Rothman DH, Xu L, & Santinelli C (2014) Hidden cycle of dissolved organic carbon in the deep ocean. *Proc. Natl. Acad. Sci. USA* 111(47):16706-16711.

22. Ciais P, *et al.* (2014) Carbon and other biogeochemical cycles. *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*, eds Stocker T, Qin D, Plattner G-K, Tignor M, Allen S, Boschung J, Nauels A, Xia Y, Bex V, & Midgley P (Cambridge University Press, New York), pp 465-570.
23. Azam F & Malfatti F (2007) Microbial structuring of marine ecosystems. *Nat. Rev. Microbiol.* 5(10):782-791.
24. Ferguson RL & Sunda WG (1984) Utilization of amino acids by planktonic marine bacteria: importance of clean technique and low substrate additions. *Limnol. Oceanogr.* 29(2):258-274.
25. Hodson RE*, et al.* (1981) Microbial uptake of dissolved organic matter in McMurdo Sound, Antarctica. *Mar. Biol.* 61(2-3):89-94.
26. Hollibaugh JT & Azam F (1983) Microbial degradation of dissolved proteins in seawater. *Limnol. Oceanogr.* 28(6):1104-1116.
27. Gifford SM, Sharma S, Booth M, & Moran MA (2013) Expression patterns reveal niche diversification in a marine microbial assemblage. *ISME J* 7(2):281-298.
28. Poretsky RS, Sun S, Mou X, & Moran MA (2010) Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environ. Microbiol.* 12(3):616-627.
29. McCarren J*, et al.* (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc. Natl. Acad. Sci. USA* 107(38):16420-16427.
30. Liu Q, Lu X, Tolar BB, Mou X, & Hollibaugh JT (2015) Concentrations, turnover rates and fluxes of polyamines in coastal waters of the South Atlantic Bight. *Biogeochemistry* 123(1-2):117-133.
31. Lidbury I, Murrell JC, & Chen Y (2014) Trimethylamine N-oxide metabolism by abundant marine heterotrophic bacteria. *Proc. Natl. Acad. Sci. USA* 111(7):2710-2715.
32. Durham BP*, et al.* (2015) Cryptic carbon and sulfur cycling between surface ocean plankton. *Proc. Natl. Acad. Sci. USA* 112(2):453-457.
33. Becker JW*, et al.* (2014) Closely related phytoplankton species produce similar suites of dissolved organic matter. *Front. Microbiol.* 5:111.
34. Bittar TB, Vieira AAH, Stubbins A, & Mopper K (2015) Competition between photochemical and biological degradation of dissolved organic matter from the cyanobacteria *Microcystis aeruginosa*. *Limnol. Oceanogr.* 60(4):1172-1194.
35. Morán XAG, Ducklow HW, & Erickson M (2013) Carbon fluxes through estuarine bacteria reflect coupling with phytoplankton. *Mar. Ecol. Prog. Ser.* 489:75-85.
36. Kaiser K & Benner R (2012) Organic matter transformations in the upper mesopelagic zone of the North Pacific: chemical composition and linkages to microbial community structure. *J Geophys Res-Oceans* 117:C01023.
37. Mopper K*, et al.* (1992) Determination of sugars in unconcentrated seawater and other natural waters by liquid chromatography and pulsed amperometric detection. *Environ. Sci. Technol.* 26(1):133-138.
38. Ankrah NY*, et al.* (2014) Phage infection of an environmentally relevant marine bacterium alters host metabolism and lysate composition. *ISME J* 8(5):1089-1100.
39. Fiore CL, Longnecker K, Kido Soule MC, & Kujawinski EB (2015) Release of ecologically relevant metabolites by the cyanobacterium *Synechococcus elongatus* CCMP 1631. *Environ. Microbiol.* 10.1111/1462-2920.12899.
40. Beier S, Rivers AR, Moran MA, & Obernosterer I (2014) The transcriptional response of prokaryotes to phytoplankton-derived dissolved organic matter in seawater. *Environ. Microbiol.* 10.1111/1462-2920.12434.
41. Raes J, Foerstner KU, & Bork P (2007) Get the most out of your metagenome: computational analysis of environmental sequence data. *Curr. Opin. Microbiol.* 10(5):490-498.

42. Keil RG & Kirchman DL (1993) Dissolved combined amino acids: chemical form and utilization by marine bacteria. *Limnol. Oceanogr.* 38(6):1256-1270.
43. Björkman KM & Karl DM (2005) Presence of dissolved nucleotides in the North Pacific Subtropical Gyre and their role in cycling of dissolved organic phosphorus. *Aquat. Microb. Ecol.* 39(2):193-203.
44. Diaz MR, Visscher PT, & Taylor BF (1992) Metabolism of dimethylsulfoniopropionate and glycine betaine by a marine bacterium. *FEMS Microbiol. Lett.* 96(1):61-65.
45. Gledhill M & Buck KN (2012) The organic complexation of iron in the marine environment: a review. *Front. Microbiol.* 3.
46. Sañudo-Wilhelmy SA*, et al.* (2012) Multiple B-vitamin depletion in large areas of the coastal ocean. *Proc. Natl. Acad. Sci. USA* 109(35):14041-14045.
47. Dyhrman ST*, et al.* (2012) The transcriptome and proteome of the diatom *Thalassiosira pseudonana* reveal a diverse phosphorus stress response. *PLoS One* 7(3):e33768.
48. Sebastian M & Ammerman JW (2009) The alkaline phosphatase PhoX is more widely distributed in marine bacteria than the classical PhoA. *ISME J* 3(5):563-572.
49. Cade-Menun BJ, Benitez-Nelson CR, Pellechia P, & Paytan A (2005) Refining ^{31}P nuclear magnetic resonance spectroscopy for marine particulate samples: storage conditions and extraction recovery. *Mar. Chem.* 97(3-4):293-306.
50. Dyhrman ST, Benitez-Nelson CR, Orchard ED, Haley ST, & Pellechia PJ (2009) A microbial source of phosphonates in oligotrophic marine systems. *Nat. Geosci.* 2(10):696-699.
51. Metcalf WW*, et al.* (2012) Synthesis of methylphosphonic acid by marine microbes: a source for methane in the aerobic ocean. *Science* 337(6098):1104-1107.
52. Van Mooy BA*, et al.* (2015) Phosphorus cycling. Major role of planktonic phosphate reduction in the marine phosphorus redox cycle. *Science* 348(6236):783-785.
53. Feingersch R*, et al.* (2012) Potential for phosphite and phosphonate utilization by *Prochlorococcus*. *ISME J* 6(4):827-834.
54. Karl DM*, et al.* (2008) Aerobic production of methane in the sea. *Nat. Geosci.* 1(7):473-478.
55. Martínez A, Ventouras LA, Wilson ST, Karl DM, & Delong EF (2013) Metatranscriptomic and functional metagenomic analysis of methylphosphonate utilization by marine bacteria. *Front. Microbiol.* 4:340.
56. Saito MA*, et al.* (2014) Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein biomarkers. *Science* 345(6201):1173-1177.
57. Glibert PM*, et al.* (2004) Evidence for dissolved organic nitrogen and phosphorus uptake during a cyanobacterial bloom in Florida Bay. *Mar. Ecol. Prog. Ser.* 280:73-83.
58. Morris JJ, Lenski RE, & Zinser ER (2012) The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *Mbio* 3(2):e00036-00012
59. Carini P, Steindler L, Beszteri S, & Giovannoni SJ (2013) Nutrient requirements for growth of the extreme oligotroph '*Candidatus* Pelagibacter ubique' HTCC1062 on a defined medium. *ISME J* 7(3):592-602.
60. Tripp HJ*, et al.* (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* 452(7188):741-744.
61. Bertrand EM & Allen AE (2012) Influence of vitamin B auxotrophy on nitrogen metabolism in eukaryotic phytoplankton. *Front. Microbiol.* 3:375.
62. Suttle CA (2007) Marine viruses - major players in the global ecosystem. *Nat. Rev. Microbiol.* 5(10):801-812.
63. Haaber J & Middelboe M (2009) Viral lysis of *Phaeocystis pouchetii*: implications for algal population dynamics and heterotrophic C, N and P cycling. *ISME J* 3(4):430-441.
64. Lønborg C, Middelboe M, & Brussaard CPD (2013) Viral lysis of *Micromonas pusilla*: impacts on dissolved organic matter production and composition. *Biogeochemistry* 116(1-3):231-240.

65. Weinbauer MG, Chen F, & Wilhelm SW (2011) Virus-mediated redistribution and partitioning of carbon in the global oceans. *Microbial carbon pump in the ocean*, eds Jiao N, Azam F, & Sanders S (Science/AAAS, Washington, DC.), pp 54-56.
66. Sherr EB & Sherr BF (1994) Bacterivory and herbivory: key roles of phagotrophic protists in pelagic food webs. *Microb. Ecol.* 28(2):223-235.
67. Sherr EB (1988) Direct use of high molecular weight polysaccharide by heterotrophic flagellates. *Nature* 335(6188):348-351.
68. Paul C, Barofsky A, Vidoudez C, & Pohnert G (2009) Diatom exudates influence metabolism and cell growth of co-cultured diatom species. *Mar. Ecol. Prog. Ser.* 389:61-70.
69. Zelezniak A*, et al.* (2015) Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc. Natl. Acad. Sci. USA* 112(20):6449-6454.
70. Johnson WM, Kido Soule MC, & Kujawinski EB (Evidence for quorum sensing and differential metabolite production by a marine bacterium in response to DMSP. *ISME J* in press.
71. Rinta‐Kanto JM*, et al.* (2011) Analysis of sulfur‐related transcription by Roseobacter communities using a taxon‐specific functional gene microarray. *Environ. Microbiol.* 13(2):453-467.
72. Amin SA*, et al.* (2015) Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature* 522(7554):98-101.
73. Persson OP*, et al.* (2009) High abundance of virulence gene homologues in marine bacteria. *Environ. Microbiol.* 11(6):1348-1357.
74. Seyedsayamdost MR, Case RJ, Kolter R, & Clardy J (2011) The Jekyll-and-Hyde chemistry of Phaeobacter gallaeciensis. *Nat. Chem.* 3(4):331-335.
75. Schaefer AL*, et al.* (2008) A new class of homoserine lactone quorum-sensing signals. *Nature* 454(7204):595-599.
76. Geng H & Belas R (2010) Molecular mechanisms underlying roseobacter-phytoplankton symbioses. *Curr. Opin. Biotechnol.* 21(3):332-338.
77. Lima-Mendez G*, et al.* (2015) Ocean plankton. Determinants of community structure in the global plankton interactome. *Science* 348(6237):1262073.
78. Stocker R (2012) Marine microbes see a sea of gradients. *Science* 338(6107):628-633.
79. Kiørboe T & Jackson GA (2001) Marine snow, organic solute plumes, and optimal chemosensory behavior of bacteria. *Limnol. Oceanogr.* 46(6):1309-1318.
80. Weitz JS*, et al.* (2015) A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes. *ISME J* 9(6):1352-1364.
81. Follows MJ, Dutkiewicz S, Grant S, & Chisholm SW (2007) Emergent biogeography of microbial communities in a model ocean. *Science* 315(5820):1843-1846.
82. Reed DC, Algar CK, Huber JA, & Dick GJ (2014) Gene-centric approach to integrating environmental genomics and biogeochemical models. *Proc. Natl. Acad. Sci. USA* 111(5):1879-1884.
83. Sowell SM*, et al.* (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* 3(1):93-105.
84. Morris RM*, et al.* (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* 4(5):673-685.
85. Shi Y, Tyson GW, Eppley JM, & DeLong EF (2011) Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. *ISME J* 5(6):999-1013.
86. Saier MH, Jr. (2000) Families of transmembrane sugar transport proteins. *Mol. Microbiol.* 35(4):699-710.
87. van der Heide T & Poolman B (2002) ABC transporters: one, two or four extracytoplasmic substrate-binding sites? *EMBO Reports* 3(10):938-943.

88. Arnosti C (2011) Microbial extracellular enzymes and the marine carbon cycle. *Ann Rev Mar Sci* 3:401-425.
89. Denger K, *et al.* (2014) Sulphoglycolysis in *Escherichia coli* K-12 closes a gap in the biogeochemical sulphur cycle. *Nature* 507(7490):114-117.
90. Mayer J, *et al.* (2010) 2, 3-Dihydroxypropane-1-sulfonate degraded by *Cupriavidus pinatubonensis* JMP134: purification of dihydroxypropanesulfonate 3-dehydrogenase. *Microbiology* 156(5):1556-1564.
91. Todd JD, *et al.* (2010) Molecular dissection of bacterial acrylate catabolism - unexpected links with dimethylsulfoniopropionate catabolism and dimethyl sulfide production. *Environ. Microbiol.* 12(2):327-343.
92. Howard EC, *et al.* (2006) Bacterial taxa that limit sulfur flux from the ocean. *Science* 314(5799):649-652.
93. Alcolombri U, *et al.* (2015) Identification of the algal dimethyl sulfide–releasing enzyme: a missing link in the marine sulfur cycle. *Science* 348(6242):1466-1469.
94. Levine NM, *et al.* (2012) Environmental, biochemical and genetic drivers of DMSP degradation and DMS production in the Sargasso Sea. *Environ. Microbiol.* 14(5):1210-1223.
95. Varaljay VA, *et al.* (2015) Single-taxon field measurements of bacterial gene regulation controlling DMSP fate. *ISME J* 9(7):1677-1686.
96. Denger K, *et al.* (2009) Bifurcated degradative pathway of 3-sulfolactate in *Roseovarius nubinhibens* ISM via sulfoacetaldehyde acetyltransferase and (S)-cysteate sulfolyase. *J. Bacteriol.* 191(18):5648-5656.
97. Lecher J, *et al.* (2009) The crystal structure of UehA in complex with ectoine-a comparison with other TRAP-T binding proteins. *J. Mol. Biol.* 389(1):58-73.
98. Chen Y, McAleer KL, & Murrell JC (2010) Monomethylamine as a nitrogen source for a nonmethylotrophic bacterium, *Agrobacterium tumefaciens*. *Appl. Environ. Microbiol.* 76(12):4102-4104.
99. Lidbury I, Kimberley G, Scanlan DJ, Murrell JC, & Chen Y (2015) Comparative genomics and mutagenesis analyses of choline metabolism in the marine *Roseobacter* clade. *Environ. Microbiol.* 10.1111/1462-2920.12943.
100. Cottrell MT, Moore JA, & Kirchman DL (1999) Chitinases from uncultured marine microorganisms. *Appl. Environ. Microbiol.* 65(6):2553-2557.
101. Wilson MC & Piel J (2013) Metagenomic approaches for exploiting uncultivated bacteria as a resource for novel biosynthetic enzymology. *Chem. Biol.* 20(5):636-647.
102. Giovannoni SJ, *et al.* (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309(5738):1242-1245.
103. Pedler BE, Aluwihare LI, & Azam F (2014) Single bacterial strain capable of significant contribution to carbon cycling in the surface ocean. *Proc. Natl. Acad. Sci. USA* 111(20):7202-7207.
104. Carlson CA, *et al.* (2002) Effect of nutrient amendments on bacterioplankton production, community structure, and DOC utilization in the northwestern Sargasso Sea. *Aquat. Microb. Ecol.* 30(1):19-36.
105. Hertkorn N, Harir M, Koch BP, Michalke B, & Schmitt-Kopplin P (2013) High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosciences* 10(3):1583-1624.
106. Jørgensen L, *et al.* (2011) Global trends in the fluorescence characteristics and distribution of marine dissolved organic matter. *Mar. Chem.* 126(1):139-148.
107. Treusch AH, *et al.* (2009) Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME J* 3(10):1148-1163.
108. Sosa OA, Gifford SM, Repeta DJ, & DeLong EF (2015) High molecular weight dissolved organic matter enrichment selects for methylotrophs in dilution to extinction cultures. *ISME J*.

109. Williams PM & Druffel ERM (1987) Radiocarbon in dissolved organic matter in the central North Pacific Ocean. *Nature* 330(6145):246-248.
110. Dittmar T & Koch BP (2006) Thermogenic organic matter dissolved in the abyssal ocean. *Mar. Chem.* 102(3):208-217.
111. Dittmar T & Paeng J (2009) A heat-induced molecular signature in marine dissolved organic matter. *Nat. Geosci.* 2(3):175-179.
112. Ziolkowski LA & Druffel ERM (2010) Aged black carbon identified in marine dissolved organic carbon. *Geophys. Res. Lett.* 37:L16601.
113. Mopper K, Kieber DJ, Stubbins A, Hansell D, & Carlson C (2015) Marine photochemistry of organic matter: processes and impacts. *Biogeochemistry of marine dissolved organic matter (second edition)*, eds Hansell DA & Carlson CA (Academic Press, Boston), pp 389-450.
114. Stubbins A, Niggemann J, & Dittmar T (2012) Photo-lability of deep ocean dissolved black carbon. *Biogeosciences* 9(5):1661-1670.
115. Medeiros PM, *et al.* (2015) Dissolved organic matter composition and photochemical transformations in the northern North Pacific Ocean. *Geophys. Res. Lett.* 42(3):863-870.
116. Lechtenfeld OJ, Hertkorn N, Shen Y, Witt M, & Benner R (2015) Marine sequestration of carbon in bacterial metabolites. *Nature Comm* 6:6711.
117. Jiao N, *et al.* (2010) Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nat. Rev. Microbiol.* 8(8):593-599.
118. Hertkorn N, *et al.* (2008) Natural organic matter and the event horizon of mass spectrometry. *Anal. Chem.* 80(23):8908-8919.
119. Arrieta JM, *et al.* (2015) Dilution limits dissolved organic carbon utilization in the deep ocean. *Science* 348(6232):331-333.
120. Dittmar T (2015) Reasons behind the long-term stability of dissolved organic matter. *The biogeochemistry of marine dissolved organic matter (second edition)*, eds Hansell DA & Carlson CA (Academic Press, Boston), pp 369-388.
121. Osterholz H, Niggemann J, Giebel H-A, Simon M, & Dittmar T (2015) Inefficient microbial production of refractory dissolved organic matter in the ocean. *Nature Comm* 6:7422.
122. Aluwihare LI, Repeta DJ, & Chen RF (1997) A major biopolymeric component to dissolved organic carbon in surface sea water. *Nature* 387(6629):166-169.
123. Carlson CA, *et al.* (2004) Interactions among dissolved organic carbon, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. *Limnol. Oceanogr.* 49(4):1073-1083.
124. Green NW, *et al.* (2014) An intercomparison of three methods for the large-scale isolation of oceanic dissolved organic matter. *Mar. Chem.* 161:14-19.
125. Arístegui J, Gasol JM, Duarte CM, & Herndl GJ (2009) Microbial oceanography of the dark ocean's pelagic realm. *Limnol. Oceanogr.* 54(5):1501-1529.
126. Beaupré SR & Aluwihare L (2010) Constraining the 2-component model of marine dissolved organic radiocarbon. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 57(16):1494-1503.
127. DeLong EF, *et al.* (2006) Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311(5760):496-503.
128. Repeta DJ & Aluwihare LI (2006) Radiocarbon analysis of neutral sugars in high-molecular-weight dissolved organic carbon: implications for organic carbon cycling. *Limnol. Oceanogr.* 51(2):1045-1053.
129. Hansman RL, *et al.* (2009) The radiocarbon signature of microorganisms in the mesopelagic ocean. *Proc. Natl. Acad. Sci. USA* 106(16):6513-6518.
130. Cho BC & Azam F (1988) Major role of bacteria in biogeochemical fluxes in the ocean's interior. *Nature* 332(6163):441-443.
131. Ingalls AE, *et al.* (2006) Quantifying archaeal community autotrophy in the mesopelagic ocean using natural radiocarbon. *Proc. Natl. Acad. Sci. USA* 103(17):6442-6447.

132. Nagata T (2000) Production mechanisms of dissolved organic matter. *Microbial ecology of the oceans,* Wiley series in ecological and applied microbiology, ed Kirchman DL (Wiley, New York), pp 121-152.
133. Swan BK*, et al.* (2011) Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science* 333(6047):1296-1300.
134. Yooseph S*, et al.* (2010) Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* 468(7320):60-66.
135. Koch BP, Witt MR, Engbrodt R, Dittmar T, & Kattner G (2005) Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochim. Cosmochim. Acta* 69(13):3299-3308.
136. Duncan KR*, et al.* (2015) Molecular networking and pattern-based genome mining improves discovery of biosynthetic gene clusters and their products from *Salinispora* species. *Chem. Biol.* 22(4):460-471.
137. Johnson CH, Ivanisevic J, Benton HP, & Siuzdak G (2015) Bioinformatics: the next frontier of metabolomics. *Anal. Chem.* 87(1):147-156.
138. Ferrera I, Sebastian M, Acinas SG, & Gasol JM (2015) Prokaryotic functional gene diversity in the sunlit ocean: stumbling in the dark. *Curr. Opin. Microbiol.* 25:33-39.
139. Lau WW & Armbrust EV (2006) Detection of glycolate oxidase gene *glcD* diversity among cultured and environmental marine bacteria. *Environ. Microbiol.* 8(10):1688-1702.
140. Teeling H*, et al.* (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336(6081):608-611.
141. da Silva RR, Dorrestein PC, & Quinn RA (2015) Illuminating the dark matter in metabolomics. *Proc. Natl. Acad. Sci. USA* 112(41):12549-12550.
142. Stedmon CA & Bro R (2008) Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. *Limnol. Oceanogr. Methods* 6(11):572-579.
143. Church MJ, Ducklow HW, & Karl DM (2002) Multi-year increases in dissolved organic matter inventories at Station ALOHA in the North Pacific Subtropical Gyre. *Limnol. Oceanogr.* 47:1-10.
144. Carlson CA, Ducklow HW, & Michaels AF (1994) Annual flux of dissolved organic carbon from the euphotic zone in the Northwestern Sargasso Sea. *Nature* 371:405-408.
145. Polovina JJ, Dunne JP, Woodworth PA, & Howell EA (2011) Projected expansion of the subtropical biome and contraction of the temperate and equatorial upwelling biomes in the North Pacific under global warming. *ICES J. Mar. Sci.* 68(6):986-995.
146. Polovina JJ, Howell EA, & Abecassis M (2008) Ocean's least productive waters are expanding. *Geophys. Res. Lett.* 35:L03618.
147. Endres S, Galgani L, Riebesell U, Schulz KG, & Engel A (2014) Stimulated bacterial growth under elevated pCO_2: results from an off-shore mesocosm study. *Plos One* 9(6):e99228.
148. Engel A*, et al.* (2014) Impact of CO_2 enrichment on organic matter dynamics during nutrient induced coastal phytoplankton blooms. *J. Plankton Res.* 36(3):641-657.
149. Chen CS, Anaya JM, Chen EY, Farr E, & Chin WC (2015) Ocean warming-acidification synergism undermines dissolved organic matter assembly. *PLoS One* 10(2):e0118300.
150. Zark M, Riebesell U, & Dittmar T (2015) Effects of ocean acidification on marine dissolved organic matter are not detectable over the succession of phytoplankton blooms. *Science Adv* 1(9):e1500531.

Figure Legends

Figure 1. Oceanic DOM is a complex mixture of molecules that are produced and consumed by billions of heterotrophic and autotrophic microbes in each liter of seawater. These heterogeneous molecules have varied reactivities towards microbial metabolism, including high reactivity (labile DOM, wide arrows) and minimal reactivity (refractory DOM, narrow arrows). Microbe-DOM interactions affect the concentration and fate of atmospheric CO_2, the accumulation of refractory carbon in the deep ocean, and flux of carbon through the ocean's food webs.

Figure 2. Significant advances that have occurred independently in three fields -- microbial ecology, geochemistry, and informatics -- have positioned oceanographers for a deeper understanding of the ocean's carbon cycle. The integration of these three fields is yielding insights into the reactions at the foundation of the global carbon cycle.