Deciphering ocean carbon in a changing world

Mary Ann Moran\textsuperscript{b,3}, Elizabeth B. Kujawinski\textsuperscript{b,4}, Aron Stubbins\textsuperscript{c,1}, Rob Fatland\textsuperscript{d,1}, Lihini I. Aluwihare\textsuperscript{e}, Alison Buchan\textsuperscript{a}, Byron C. Crump\textsuperscript{g}, Pieter C. Dorrestein\textsuperscript{h,i,j}, Sonya T. Dyhrman\textsuperscript{h,j,l}, Nancy J. Hess\textsuperscript{m}, Bill Howe\textsuperscript{n}, Krista Longnecker\textsuperscript{p}, Patricia M. Medeiros\textsuperscript{q}, Jutta Niggemann\textsuperscript{k}, Ingrid Obernosterer\textsuperscript{p}, Daniel J. Repeta\textsuperscript{q}, and Jacob R. Waldbauer\textsuperscript{q}

Edited by David M. Karl, University of Hawaii, Honolulu, HI, and approved February 11, 2016 (received for review October 21, 2015)

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\textsubscript{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\textsubscript{2} in the oceans, 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.

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\textsubscript{2} reservoir (2) (Fig. 1).

Flux of carbon through the marine DOM pool is mediated largely by microbial activity. However, the intertwined relationships between the molecules making up the DOM pool and the ocean microbes that process them remain poorly characterized. The complexity of each has defied easy characterization, and fundamental interactions have been necessarily oversimplified 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
loosely corresponds to a depth gradient. In the surface ocean, the categories occur throughout the ocean, their relative importance least biologically reactive and circulates through the major ocean for weeks to years. Production (Fig. 1).

Molecules that are consumed by microbes within hours or days of production are typically degraded extracellularly into oligomers or monomers before 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 monocarboxylic and dicarboxylic acids (27, 28), glycerols and fatty acids (27, 29), and the nitrogen-containing metabolites taurine, choline, sarcosine, polyamines, methyamines, 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: Photosynthetic release 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 osmylates (44), siderophores (45), vitamins (46), and primary metabolites. 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 gigaton or 1 × 10^15 g). 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 before 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 monocarboxylic and dicarboxylic acids (27, 28), glycerols and fatty acids (27, 29), and the nitrogen-containing metabolites taurine, choline, sarcosine, polyamines, methyamines, 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: Photosynthetic release 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 osmylates (44), siderophores (45), vitamins (46), and primary metabolites.
Advances in understanding the fate of these diverse components of DOM have occurred despite 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 longstanding 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 by both 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 cofactors 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 Pelagibacteriales 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 B12 and B2 (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, whereas 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 noninfected cells (38). New categories of nonpredatory 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 cooccurrences 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 with 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? Although there is currently no way to know the full biochemical diversity behind microbial processing of DOM, headway is being made in understanding the fate of these diverse components of DOM that were previously not measurable (39, 46).
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).

However, assigning an exact function to microbial genes remains a stubborn obstacle 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 to use model microbial systems amenable to experimental manipulation and genetic modification. For example, the first marine bacterial gene mediating catabolism of dimethylsulfonioacetate (DMSP), a compound known for 50 y to be critical in marine bacterial gene mediating catabolism of dimethylsulfonioacetate (DMSP), a compound known for 50 y to be critical in marine bacterial carbon cycling, 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 to both 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, or 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 Semilabile DOM Accumulate in the Surface Ocean?

Semilabile 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 by substrates and nutrients remains a mystery. As a substantial and temporally stable component of DOM, the semilabile pool affects the overall rate of carbon turnover in the oceans (Fig. 1). Therefore, 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 covary with semilabile DOM and depth in the ocean, and thereby helping in 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 Pelagibacteraceae dominate open ocean gyres where semilabile 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 Pelagibacteraceae, showed that, although this one strain can degrade the labile fraction of marine DOM in a period of days, an amount of DOM equivalent to the semilabile pool remained untouched. Instead, the full microbial community was needed to degrade the semilabile DOM (103). Earlier studies showed that the addition of both labile DOM and inorganic nutrients is needed to degrade semilabile DOM (104), signifying a complex relationship between DOM accumulation, microbial diversity, and the availability of nutrients and cometabolites.

The chemical and optical signatures of seawater also differ from the deep ocean background in locations where semilabile 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 semilabile DOM relative to the signatures of deep ocean DOM (105, 106). Nevertheless, extraction protocols are insufficient and chemical understanding is too limited to physically isolate semilabile 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. However, this is the repository for over 70% of the carbon sequestered in DOM and is 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 y (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 semilabile (radiocarbon enriched) and refractory (radiocarbon deplete) pools within the DOM, with the most deplete fraction having a radiocarbon age of ~12,000 y (21).

That these energy-rich molecules exist for millennia in the deep ocean is paradoxical and 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 of 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, although it is home to the largest reservoir of refractory DOM, it also harbors labile and semilabile 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).
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 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 winter time area between 1998 and 2006 (145, 146), may therefore increase net oceanic DOM production. However, experimental studies suggest that rising temperatures and ocean acidification will increase bacterial DOM consumption (147, 148), whereas 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 square kilometers 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.

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. However, 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 tools for 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, nontargeted 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). ‘Oms 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, and 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 crowdsourcing to solve common problems in compound identification (gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp). 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).

Acknowledgments
We thank Jack Cook for graphics expertise. 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 OCE13566010, OCE1154320, and OCE1356890, and by Gordon and Betty Moore Foundation Grant 3304.
Geng H, Belas R (2010) Molecular mechanisms underlying


Denger K, et al. (2014) Sulphoglycolysis in


Johnson WM, Kido Soule MC, Kujawinski EB (2016) Evidence for quorum sensing and differential metabolite production by a marine bacterium in response to


Rinta-Kanto JM, et al. (2011) Analysis of sulfur-related transcription by

Johnson (2009) The crystal structure of UehA in complex with ectoine-A comparison with other TRAP-T binding proteins.


Johnson (2009) Bifurcated degradative pathway of 3-sulfolactate in


