Phosphonate production by marine microbes: Exploring new sources and potential function

Marianne Acker[a,b,1], Shane L. Hogle[c,d,1] (iD), Paul M. Berube[c] (iD), Thomas Hackl[c], Allison Coe[c] (iD), Ramunas Stepanauskas[e] (iD), Sallie W. Chisholm[c,f] (iD), and Daniel J. Repeta[b,2] (iD)

[a]Massachusetts Institute of Technology-Woods Hole Oceanographic Institution Joint Program in Oceanography/Applied Ocean Science and Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA 02543; [b]Department of Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543; [c]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139; [d]Department of Biology, University of Turku, Turku 20500, Finland; [e]Single Cell Genomics Center, Bigelow Laboratory for Ocean Sciences, East Boothbay, ME 04544; and [f]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Phosphonates are organophosphorus metabolites with a characteristic C-P bond. They are ubiquitous in the marine environment, their degradation broadly supports ecosystem productivity, and they are key components of the marine phosphorus (P) cycle. However, the microbial producers that sustain the large oceanic inventory of phosphonates as well as the physiological and ecological roles of phosphonates are enigmatic. Here, we show that phosphonate synthesis genes are rare but widely distributed among diverse bacteria and archaea, including *Prochlorococcus* and SAR11, the two major groups of bacteria in the ocean. In addition, we show that *Prochlorococcus* can allocate over 40% of its total cellular P-quota toward phosphonate production. However, we find no evidence that *Prochlorococcus* uses phosphonates for surplus P storage, and nearly all producer genomes lack the genes necessary to degrade and assimilate phosphonates. Instead, we postulate that phosphonates are associated with cell-surface glycoproteins, suggesting that phosphonates mediate ecological interactions between the cell and its surrounding environment. Our findings indicate that the oligotrophic surface ocean phosphonate pool is sustained by a relatively small fraction of the bacterioplankton cells allocating a significant portion of their P quotas toward secondary metabolism and away from growth and reproduction.

phosphonate | *Prochlorococcus* | marine | biogeochemistry | phosphorus

Phosphorus (P) is an essential nutrient for all life. Phytoplankton and bacterioplankton (bacteria and archaea) in the surface ocean use P primarily for growth, reproduction, and as an energy currency such that almost all cellular P is associated with rapidly cycled compounds, including nucleic acids, ATP, and phospholipids (1). For example, in the North Pacific subtropical gyre, RNA synthesis accounts for over half of phosphate uptake in phytoplankton communities (2). In times of P scarcity, some marine microbes can modulate their cellular P requirements by recycling RNA from ribosomes (3), substituting phospholipids with other non-P lipids (4), or drawing upon internal stores of P accumulated during periods of P abundance (5, 6). The balance between microbial, physiological processes that require P and those that conserve or internally recycle P has significant downstream consequences for ecological stoichiometry and the marine P biogeochemical cycle (7).

Nevertheless, in nutrient-impoverished midocean gyres, microbial demand for P is often so high that environmental phosphate concentrations are consumed to low-nanomolar levels. Under these conditions, up to half of the microbial P demand is met through the uptake and metabolism of P-containing dissolved organic matter (8). Dissolved organic P (DOP) is a complex, poorly characterized mixture of high– and low–molecular weight (HMW and LMW) phosphate and phosphonate esters released via exudation and cell death. Marine DOP consists primarily of monomeric and polymeric phosphate esters originating from well-known primary metabolites (e.g.,

nucleotides, nucleic acids, phospholipids, phosphoglycans, phosphoproteins, and vitamins) that are synthesized by all marine microbes. The presence of phosphate esters in DOP is easily explained.

However, a large fraction of HMWDOP, between 20 and 25%, occurs as phosphonates (9)—an enigmatic class of compounds with a direct carbon to P (C-P) bond—that are not products of the primary cellular metabolism (Fig. 1) (10). Phosphonates were initially believed to be rare and unusual compounds with little environmental relevance (11). However, recent studies have indicated that many marine microbes have the potential to degrade and use phosphonates as a nutritional P source (Fig. 1) (12–14). This has been experimentally confirmed (15–17), showing that degradation of phosphonates in HMWDOP is an important source of marine methane, thereby contributing to the supersaturation of methane in oxic oligotrophic waters (18)—the so-called "marine methane paradox" (19).

Although it is now recognized that phosphonates support a significant fraction of microbial P demand in the ocean, only two bacterioplankton species have been experimentally confirmed as phosphonate producers: *Trichodesmium erythraeum*, a

Significance

Phosphonates are a class of phosphorus metabolites characterized by a highly stable C-P bond. Phosphonates accumulate to high concentrations in seawater, fuel a large fraction of marine methane production, and serve as a source of phosphorus to microbes inhabiting nutrient-limited regions of the oligotrophic ocean. Here, we show that 15% of all bacterioplankton in the surface ocean have genes phosphonate synthesis and that most belong to the abundant groups *Prochlorococcus* and SAR11. Genomic and chemical evidence suggests that phosphonates are incorporated into cell-surface phosphonoglycoproteins that may act to mitigate cell mortality by grazing and viral lysis. These results underscore the large global biogeochemical impact of relatively rare but highly expressed traits in numerically abundant groups of marine bacteria.
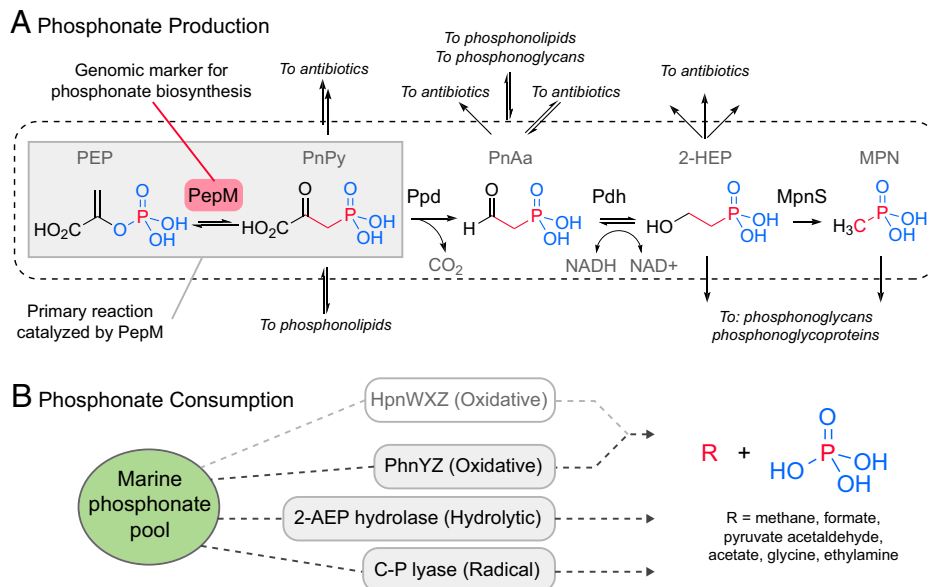
ENVIRONMENTAL SCIENCES

Fig. 1. Phosphonate production and consumption in the marine environment. (*A*) This schematic emphasizes reaction pathways (e.g., to methylphosphonate [Mpn]) (18) that are most relevant for planktonic ecosystems in the oligotrophic surface ocean. Branch points to additional natural products are indicated where possible. For a comprehensive overview please refer to ref. 10. The isomerization of phosphoenolpyruvate (PEP) via PEP mutase (PepM, red and gray boxes) produces the phosphonopyruvate (PnPy) precursor from which effectively all known phosphonate compounds are derived. The *pepM* gene is a robust genomic marker for phosphonate production potential due to this functional conservation. Mpn, a well-known substrate for aerobic marine methane production (18), is produced by 1) the decarboxylation of PnPy to phosphonoacetaldehyde (PnAa) via phosphonopyruvate decarboxylase (Ppd), 2) the dehydrogenation of PnAa to 2-hydroxyethylphosphonic acid (2-HEP) via phosphonoacetaldehyde dehydrogenase (Pdh), and finally 3) transformation of 2-HEP by Mpn synthase (MpnS). (*B*) Phosphonate degradation via cleavage of the C-P bond proceeds through at least three mechanisms: hydrolytic (2-AEP, named for the representative 2-aminoethylphosphonate degradation pathway via PhnWX), radical (C-P lyase), and oxidative (PhnYZ). The proposed HpnWXZ pathway (greyed) is likely oxidative but is rare in the genomes we examined.

nitrogen-fixing cyanobacterium (20), and *Nitrosopumilus mariti-mus* from the Marine Group I (MGI) Thaumarchaeota (21). These two groups by themselves can only account for a small fraction of the total inventory of phosphonates observed in surface ocean DOP (*SI Appendix, SI Results*) (20).

Most of our understanding of the distribution of the phosphonate synthesis trait in the wild is derived from genomics approaches using phosphonate biosynthesis genes as biomarkers (21–23). The enzyme phosphonenolpyruvate mutase (PepM) is required to form the initial C-P bond in the phosphonopyruvate precursor from which all other known phosphonate biosynthetic products are derived (Fig. 1). This makes the *pepM* gene a key biomarker for phosphonate biosynthetic potential. Metagenomic studies have estimated that 8 to 16% of all surface bacterioplankton may be phosphonate producers, including the abundant marine groups *Prochlorococcus* and SAR11 (21, 22). However, none of these organisms have been experimentally shown to produce phosphonate, and little is known about the functional roles that phosphonates play in *Prochlorococcus* and SAR11 ecology (10).

Here, we first expand on earlier metagenomic studies and incorporate thousands of randomly sampled single-cell genomes from across the tropical and subtropical surface ocean to identify potential phosphonate producers and to quantify their potential contribution to the phosphonate pool. The single-cell genomes allowed us to determine whether individual cells are capable of both producing and consuming phosphonates. The expanded dataset also allowed us to explore the frequencies of phosphonate production and consumption traits among closely related groups and identify associations of these traits with environmental variables. Further, we identified a cultured strain of *Prochlorococcus* that has the full phosphonate biosynthetic pathway and demonstrate that it produces high–molecular weight phosphonate compounds. Finally, we

identified the macromolecular cellular components with which the phosphonates are associated giving us clues as to their functional roles in ocean ecosystems.

Results and Discussion

Phosphonate-Cycling Genes Are Common in SAR11 and *Prochlorococcus*. How widespread is the putative ability to produce and consume phosphonates among diverse marine microbes? To address this question, we used a genome-resolved approach with thousands of randomly sampled single-cell amplified genomes (SAGs) from the Global Ocean Reference Genome (GORG-Tropics) dataset (24), which contained cells from the tropical and subtropical surface ocean. The sequenced cells in this database were randomly selected, thus minimizing bias. Assembly ambiguities are further avoided, because the genomic information is known to come from a single cell. Putative phosphonate producers—as defined by the presence of the *pepM* gene (Fig. 1)—were found in 15% of the bacterioplankton, belonging to diverse taxa, in this dataset (*SI Appendix, SI Results* and Fig. S1 and Datasets S1 and S2). A total 9% of the cells in a single deeply sequenced sample from the North Atlantic (GORG-BATS248) were producers (*SI Appendix*, Table S1). The two most-numerically abundant and cosmopolitan marine groups in the tropical surface ocean—SAR11 and *Prochlorococcus*—have the potential to be significant phosphonate producers, as the trait is found in 18% and 6% of their cells, respectively. Together, SAR11 and *Prochlorococcus* account for over half of the cells with this trait in the GORG dataset (Fig. 2 *A* and *B*). SAR11 are small, chemoheterotrophic Alphaproteobacteria estimated to constitute up to half of the total plankton cells in the surface ocean (25), while *Prochlorococcus* numerically dominates unicellular, photoautotrophic picocyanobacteria in the euphotic zone of subtropical and tropical oligotrophic areas (26).
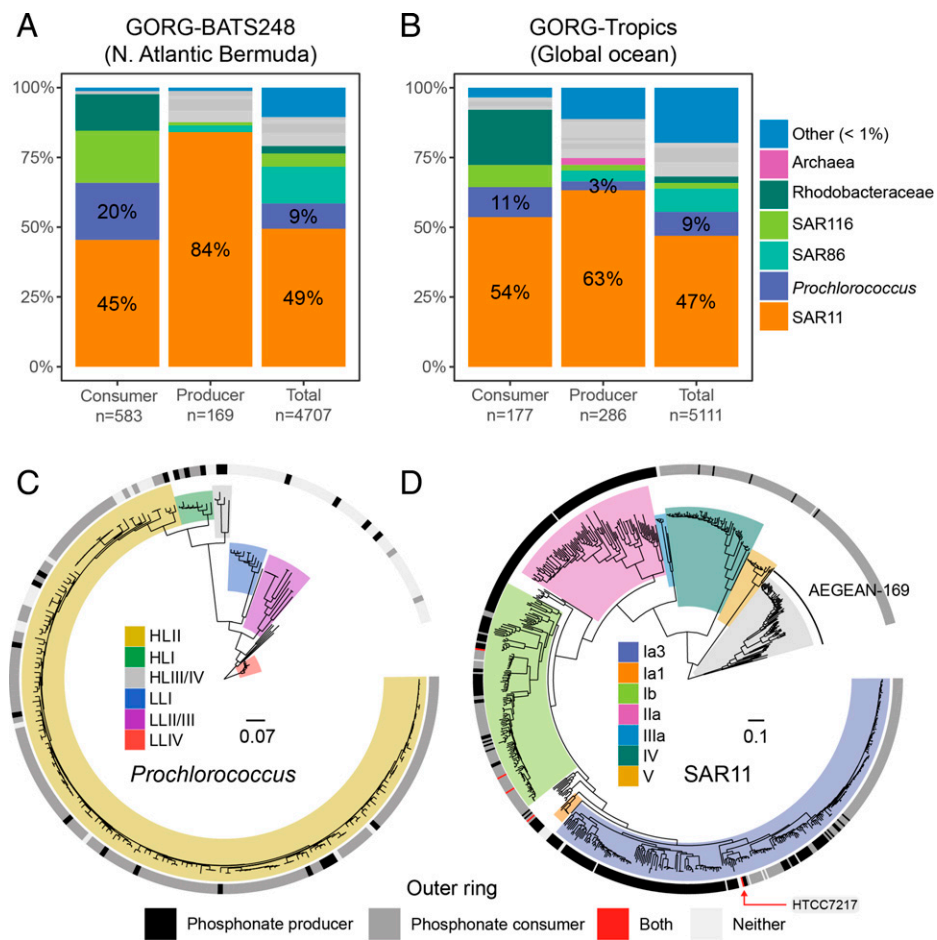
Fig. 2. SAR11 and *Prochlorococcus* potential for phosphonate production and consumption in the surface ocean. Taxonomic assignments (as percent of total cells) of phosphonate consumers, producers, and all (total) genomes from the following: (*A*) Bermuda in the North Atlantic Ocean (GORG-BATS248) and (*B*) the global tropical ocean (GORG-Tropics). The number of genomes in each category are listed on the *x*-axis. Taxonomic groups with abundance less than 1% of the total are collapsed. Gray bars represent nonfocal taxonomic groups (*SI Appendix*, Fig. S1). Here, GORG-Tropics contains a random selection of genomes from GORG-BATS248 subsampled to the median per-sample abundance in GORG-Tropics. Genome phylogenies of (*C*) *Prochlorococ-cus* (n = 235) and (*D*) SAR11 (n = 695) single-cell genomes from GORG-Tropics and isolate genomes. Dominant clades/ecotypes from each group are highlighted. The ring outside each tree displays whether a genome is a phosphonate producer (black), a phosphonate consumer (gray), or both a producer and consumer (red). Isolate genomes with neither pathway are shown in white. For reference, HTCC7217 from the SAR11 clade 1a3 is both a producer and consumer. *Prochlorococcus* is rooted on the LLIV clade and SAR11 is rooted on the AEGEAN-169 group. Scale bars represent amino acid substitution per site.

We next asked how abundant phosphonate degradation pathways were in the GORG-Tropics dataset. Approximately 10% of all bacterioplankton in the tropical surface ocean encode genes for phosphonate consumption—defined by the presence of at least one of four distinct degradation pathways (Fig. 1 and *SI Appendix*, Table S1). A total 31% of the cells in GORG-BATS248 (i.e., at a single station in the Atlantic) had the genomic potential to degrade phosphonates, consistent with relatively high North Atlantic consumer abundances inferred from metagenomic studies (27). Overall, however, the proportions of putative producers and consumers are similar across the entire tropical surface ocean (15 and 10%, respectively).

Phosphonate Production and Consumption Traits Are Rarely Found in the Same Cell. Both putative phosphonate producers and consumers can be found within closely related taxonomic groups in the GORG dataset (e.g., the *Prochlorococcus* genus contains both). Therefore, we next asked whether phosphonate production and consumption traits were partitioned at the level of clades within these groups or at the level of single cells. Neither production nor consumption traits were correlated with the *Prochloro-coccus* phylogeny, indicating that most *Prochlorococcus* clades

contain both phosphonate producers and consumers (Fig. 2*C*). In SAR11, the correlation between traits and phylogeny was stronger, particularly for clades IIa, IV, and V—the former one being producers and the latter clades mostly consumers (Fig. 2*D*). However, the highly abundant SAR11 surface clades Ia and Ib contained both producers and consumers in roughly equal proportions. Remarkably, only 0.7% (9/1,234) of all genomes contained both production and consumption traits (i.e., in the SAR11, SAR116, and Roseobacter groups). The presence of both phosphonate consumption and production traits may be advantageous for some bacterial lifestyles (e.g., rare facultative oligotrophs that can also quickly respond to resource pulses). However, the rarity of this trait combination in the wild suggests that it is prohibitively costly for most bacteria.

We also examined tens of thousands of other microbial genomes from marine, freshwater, terrestrial, and host-associated habitats (*Materials and Methods*) to determine whether the pattern of production and consumption exclusivity held over a wider taxonomic and habitat breadth. As with the GORG database, less than 1% of these genomes were both producers and consumers (*SI Appendix*, Fig. S2). All genomes with both traits had high-scoring hits to phosphonate production

Acker et al.
Phosphonate production by marine microbes: Exploring new sources and potential function

PNAS | 3 of 12
https://doi.org/10.1073/pnas.2113386119

ENVIRONMENTAL SCIENCES

and consumption genes. Although we identified the SAR11 strain HTCC7217 and two marine Roseobacter isolates (*Actibacterium atlanticum* 22II-S11-z10 and *Pseudooceanicola nanhaiensis* DSM 18065) with both a C-P lyase gene cluster and phosphonate biosynthesis genes, the vast majority of genomes with both production and consumption traits in this dataset were nonmarine (Dataset S3). For example, almost half were from the Burkholderiaceae family of the Betaproteobacteria (*SI Appendix*, Fig. S2), which included many human pathogens. Thus, the general mutual exclusivity between phosphonate biosynthesis and catabolism traits is not unique to the marine environment. In surface ocean clades of *Prochlorococcus* and SAR11, this mutual exclusivity occurs at the species/strain level, which implies strong functional differentiation between closely related cells (Fig. 2 *C* and *D*) and fine-scale niche partitioning with regard to phosphonate function and may reflect functional incompatibility or ecological/evolutionary tradeoffs between biosynthesis and catabolism.

Phosphonate Production Traits Are Horizontally Transferred. Because phosphonate production and consumption traits are found in a subset of the genomes from each taxonomic group (*SI Appendix*, Table S1), they are part of their flexible genome—as has already been demonstrated for both SAR11 and *Prochlorococcus* (28). Horizontal gene exchange is a key mechanism maintaining genes at medium-low frequencies within the flexible genome of microbial populations (29); thus, we looked for its signatures with regard to phosphonate production genes. We found that their gene-flow patterns do not follow a tree-like pattern and are better explained by horizontal exchange and gene loss. The topological distance between a species phylogeny and a *pepM* phylogeny approaches values for random tree sets and is over twice the distance between the species tree and trees from conserved metabolic gene families (*SI Appendix*, Fig. S3). Thus, there is a poor correlation between species trees, inferred from conserved marker genes, and trees constructed from the *pepM* gene—a hallmark of horizontal gene transfer (30). Furthermore, nearly all phosphonate biosynthesis gene cassettes are located within "genomic islands" in *Prochlorococcus* (*SI Appendix*, Fig. S4), which contain the majority of laterally transferred genes in this genus (31, 32). There is also a transposase—an important molecular mechanism of horizontal gene transfer in bacteria (33)—four genes downstream of *pepM* in the *Prochlorococcus* SB genome (*SI Appendix*, Fig. S5). Overall, these findings suggest that the phosphonate production trait has been frequently exchanged among marine microbial communities.

Oceanographic Trends in Phosphonate-Producing Microbes. How do the global distributions of putative phosphonate producers and consumers relate to hydrographic and biogeochemical patterns? Results from GORG-Tropics revealed two significant trends: 1) the abundance of putative phosphonate producers significantly increases with depth and is slightly higher in the Atlantic Ocean (Fig. 3*A* and *SI Appendix*, Fig. S6) and 2) putative phosphonate consumers, including both SAR11 and *Prochlorococcus*, are significantly more abundant in the Atlantic than in the Pacific Ocean (Fig. 3*B* and *SI Appendix*, Fig. S6), which is consistent with past reports (27, 34). We found very high consumer abundance in four samples from the northwestern edge of the Sargasso Sea that were collected during late autumn. However, even after excluding these four samples the Pacific/Atlantic difference was still significant ($P = 0.02$). We then greatly increased the temporal and spatial resolution of our analysis by using nearly 700 metagenomes from GEOTRACES, Tara Oceans, and two long-term ocean time series sites (35–37) in order to look for other relationships between putative phosphonate producers and environmental variables. Based on metagenome read recruitment to *pepM* and conserved marker

genes (*Materials and Methods*), we estimate that 15% of bacterioplankton genomes are putative phosphonate producers (Fig. 3*C*). A total 6% of *Prochlorococcus* carry the trait, as do 10% of SAR11 (Fig. 3*C*), corroborating the numbers we found in the GORG-Tropics SAGs (*SI Appendix*, Table S1). Likewise, the dominant trends we observed in GORG-Tropics were clearly recovered by the metagenomes. There was no statistically significant difference in producer abundance between ocean basins (*SI Appendix*, Fig. S7A), and the abundance of cells carrying phosphonate production genes increased with depth, with a rapid increase starting near 100 m (Fig. 3*D* and *SI Appendix*, Fig. S8). The agreement between the metagenome and single-cell genome–resolved approaches highlights the robustness of our findings.

To better understand what drives the distribution of phosphonate producers in the sunlit ocean, we next incorporated these metagenomic datasets along with chemical, biological, and hydrographic data into predictive models. The models suggest that on average, *Prochlorococcus* phosphonate producers are most abundant where inorganic phosphate and DOP concentrations are highest (*SI Appendix*, Figs. S7 and S8). For example, the North Atlantic and Mediterranean had the lowest median abundance of *Prochlorococcus* producers. In GORG-BATS (Bermuda Atlantic Time-series Study), we detected no *Prochlorococcus* phosphonate producer genomes even though the sample was deeply sequenced (>6,000 genomes). We interpret this as reflecting potential selection against the phosphonate biosynthesis trait in *Prochlorococcus* populations from consistently P-limited regions due to the additional P cost of phosphonate biosynthesis. The relative abundances of SAR11 producers and total producers were shaped by many of the same factors consistent with SAR11 being the dominant phosphonate producer in GORG-tropics and GORG-BATS. We found that the greatest amount of phosphonate producer variation is explained by multiple depth-dependent biotic factors (*SI Appendix*, Figs. S7 and S8). Overall, the models predict that phosphonate producers are most abundant in the lower reaches of the euphotic zone and upper reaches of the mesopelagic. Future chemical studies targeting the 100- to 300-m depth range will hopefully provide greater understanding of phosphonate production in the upper ocean.

We next looked for temporal patterns from metagenomes collected from the Hawaii Ocean Time-series (HOT) and the BATS, two long-running time series representative of the North Pacific and North Atlantic subtropical gyres (38, 39). BATS has a lower-total dissolved P inventory than HOT, which has been linked to variation in P-related gene content in BATS *Prochlorococcus* populations (28). Putative phosphonate producer abundance peaked during late winter and early spring at BATS, coinciding with wind-driven deep mixing (*SI Appendix*, Fig. S9C), which increases the supply of inorganic phosphate to the surface waters (40) and may also recruit phosphonate producers from deeper layers. Similarly, phosphonate-producing SAR11 ecotypes IIa and IIb are abundant in the winter/spring (41), while phosphonate-consuming SAR11 ecotypes IV and V dominate in the highly stratified summer months when inorganic phosphate concentrations are low (Fig. 2*D*). Likewise, putative *Prochlorococcus* producers declined to negligible levels in the summer. This implies that the phosphonate production trait occurs at the highest frequency when inorganic phosphate concentrations are relatively high, possibly reflecting a high cost of production.

Based on the seasonal results at BATS and our intuition that phosphonate production might increase overall cellular P demand, we asked whether the phosphonate production trait was less abundant in low P regions of the ocean or whether the trait was associated with other oceanographic features. We found a small but significant positive relationship between inorganic phosphate and DOP concentrations and the abundance of *Prochlorococcus* phosphonate producers (*SI Appendix*, Figs.
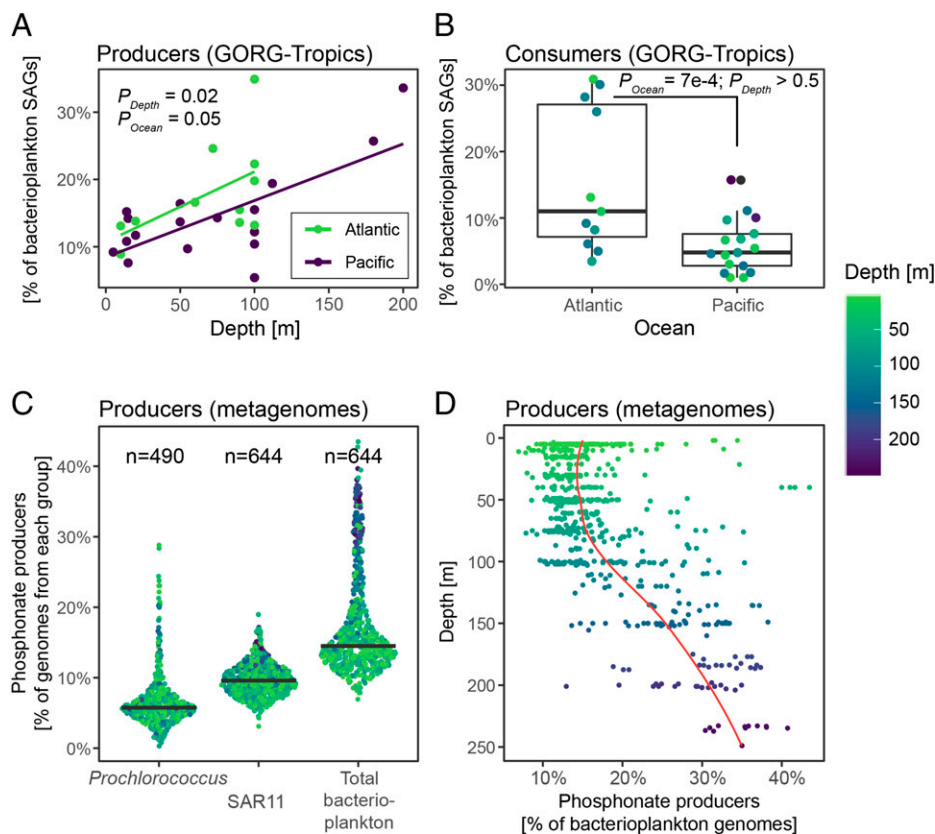
Fig. 3. Oceanographic drivers of phosphonate production. (*A*) The proportion of phosphonate producers significantly increases with depth in the 28 GORG-Tropics samples (beta-binomial regression; Est = 0.004, Err = 0.001, t = 2.604, and P = 0.02; link = logit; log L = −78.179, df = 4, and resid df = 24). Producers are slightly more common in the Atlantic than the Pacific (beta-binomial regression; Est = −0.30, Err = 0.14, t = −2.08, and P = 0.05; link = logit; log L = −78.179, df = 4, and resid df = 24). Each point is a sample (median SAGs per sample = 241). (*B*) Phosphonate consumers in the Atlantic and Pacific oceans as percent of all SAGs (Est = −1.04, Err = 0.27, t = −3.81, and P = 7 × 10⁻⁴; link = logit; log L = −80.425, df = 4, and resid df = 24). Proportions (%) are total producers or consumers divided by the total number GORG assemblies and are corrected using the estimated sequence recovery from assemblies (*Materials and Methods*). (*C*) Percentage of phosphonate producers in the global (sub)tropical surface ocean estimated from BioGEOTRACES and Tara Oceans metagenomes. Total bacterioplankton category includes both Bacteria and Archaea. Black line is the median value for n metagenome samples for each taxonomic group. (*D*) The relationship between depth and percent of all bacterioplankton with the phosphonate production trait. The red line is a simple Loess regression fit to the data.

This is consistent with the seasonal variation of phosphonate producers at BATS and implies that relatively high inorganic phosphate and DOP conditions support more cells with genes encoding for phosphonate production. However, we did not observe a relationship between inorganic phosphate or DOP and putative SAR11 or total community producers. Thus, the relationship between community-averaged phosphonate production and oceanographic P inventories is unclear. For *Prochlorococcus*, at least, higher-total dissolved P concentrations seem to be associated with a higher fraction of cells with the genetic potential for phosphonate production.

Exploring the depth dependence of phosphonate producers yielded a few additional insights. In samples shallower than 100 m, a median of 15% of bacterioplankton have phosphonate biosynthesis genes. That proportion steadily increases with depth to nearly 30 to 40% at 200 m (Fig. 3D). The 100-m inflection coincides with the nutrient-driven "genomic transition zone" (i.e., a region where bacteria and archaea tend to have larger genomes with higher GC content and proteins with higher N/C ratios) (42). Again, nutrient availability, which increases with depth, is an important factor in the relative abundance of putative phosphonate producers. Finally, the proportion of putative phosphonate producers in the upper 100 m of the tropical ocean is relatively modest (<15%). Since phosphonates appear to be relatively rapidly cycled (18) and constitute a significant fraction of the DOP, this implies that cellular production rates would need to be relatively high to account for the large global inventory of phosphonates in marine dissolved organic matter.

Phosphonate Production in *Prochlorococcus* SB. To better understand how much cellular P is allocated to phosphonates and with which macromolecules they are associated, we sought to examine phosphonate production by an abundant, widespread, and experimentally tractable marine microbe. Only one cultured isolate meets these criteria: *Prochlorococcus* SB. Since there is no genetic system for knockout mutants in *Prochlorococcus*, we used *Prochlorococcus* MIT9301, a closely related strain lacking the phosphonate biosynthesis cluster, for comparison. Using ³¹P NMR spectroscopy (³¹P-NMR), we found that under P-replete conditions *Prochlorococcus* MIT9301 allocates all its P to phosphate and pyrophosphate esters (−12 to 12 ppm; Fig. 4*B*). In contrast, *Prochlorococcus* SB—the strain with the phosphonate biosynthetic pathway—produces both phosphate and pyrophosphate esters as well as phosphonate esters (43) (18 to 27 ppm; Fig. 4*B*). Some cyclic phosphates, specifically 1, 3, 2-dioxaphospholanes such as 2′, 3′-cyclic nucleotide monophosphates, have chemical shifts in the 14- to 20-ppm spectral region and could be misassigned as phosphonates based on their ³¹P-NMR spectra alone. We therefore treated
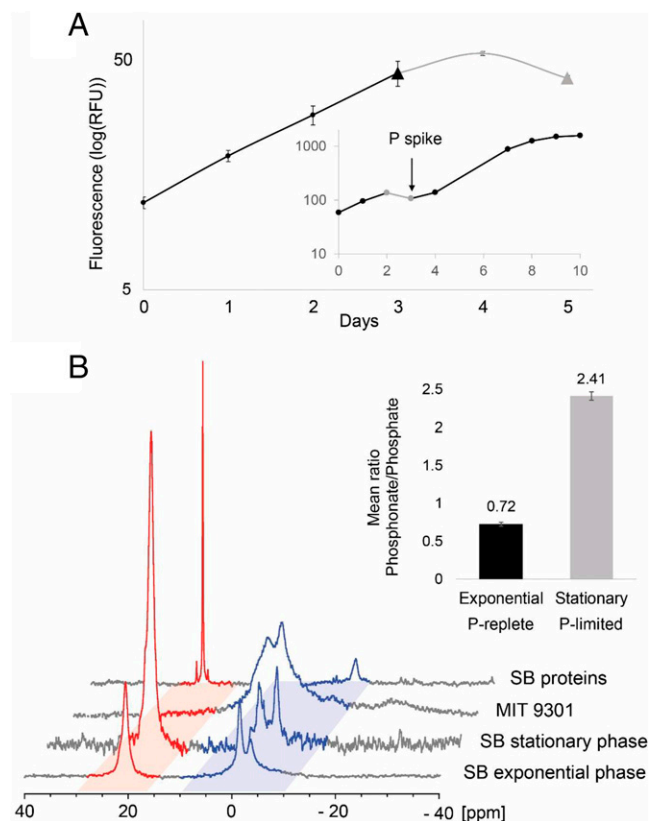
Acker et al.
Phosphonate production by marine microbes: Exploring new sources and potential function

PNAS | 5 of 12
https://doi.org/10.1073/pnas.2113386119

Fig. 4. Phosphonate production by cultured *Prochlorococcus* SB. (*A*) *Prochlorococcus* SB growth curve in medium with N/*P* = 350/1 showing exponential phase (P-replete; black line and symbols) and stationary phase (P-starved; gray line and symbols) due to P-starvation. Triangle data points correspond to the days the cells were harvested in the different growth phases. Error bars are calculated based on the SD between the biological duplicates. The *Inset* represents the *Prochlorococcus* SB growth curve in medium with N/*P* = 350/1 in which inorganic phosphate was added on Day 3 to reach N/*P* = 16/1. (*B*) ^{31}P-NMR spectra of exponential (P-replete) and stationary (P-starved) phase *Prochlorococcus* SB whole cells, P-replete *Prochlorococcus* MIT9301 whole cells, and the insoluble protein fraction of P-replete *Prochlorococcus* SB. The phosphonate and phosphate regions of the spectra are indicated in red and blue, respectively. While *Prochlorococcus* SB produces phosphonate and doubles its relative phosphonate content in P-starved stationary phase, the negative control, MIT9301 only produces phosphates. The histogram (*Inset*) displays the mean phosphonate/phosphate ratios for *Prochlorococcus* SB cells harvested in exponential (black) and stationary (gray) phase calculated by integrating the phosphonate and phosphate peaks in *Prochlorococcus* SB whole-cell ^{31}P-NMR spectra obtained for the duplicates in each growth phase. Error bars correspond to the SD of the biological replicate.

Prochlorococcus SB with 0.1 N NaOH using conditions that are known to degrade 2',3'-cyclic nucleotides. ^{31}P-NMR spectra collected before and after base-catalyzed hydrolysis showed no measurable difference in the integrals of the phosphonate/phosphate ester regions, indicating that 2', 3'-cyclic nucleotide monophosphates and similar 1, 3, 2-dioxaphospholanes do not contribute significantly to P signal in the 18- to 27-ppm region of the ^{31}P-NMR spectra (*SI Appendix, SI Results* and Fig. S12). Integration of the phosphate and phosphonate ester regions of the NMR spectrum indicated that under nutrient replete conditions, *Prochlorococcus* SB cells allocate ~40% of their cellular P to phosphonate production (Fig. 4*B* and *SI Appendix*, Table S3).

Phosphonates and P Storage. The physiological and ecological roles of phosphonates are poorly understood. However, the presence of *pepM* among diverse, genome-streamlined bacteria

and the high relative abundance of cellular phosphonates in *Prochlorococcus* SB suggest that phosphonates serve an essential function for microbes inhabiting oligotrophic marine waters. In highly stratified oligotrophic waters, nutrient supply to phytoplankton is episodic, driven by mixing events that bring nutrient-rich waters from below the surface into the euphotic zone. Some microbes take up excess P during periods of high nutrient concentrations and sequester them for later use (5, 6); thus, one function that has been suggested for phosphonates is that they could serve as an intracellular P-storage reservoir (44). The *Prochlorococcus* SB genome, however, lacks phosphonate degradation pathways, including C-P lyase, phosphonate hydrolytic pathways (10), and PhnYZ, an oxidative pathway recently discovered in another strain of *Prochlorococcus* (27).

Recognizing that *Prochlorococcus* SB might use an uncharacterized pathway to repurpose P from phosphonates, we tested the P-storage hypothesis in this strain by comparing phosphonate production in cultures grown under P-starved and P-replete conditions (Fig. 4*A*), expecting that if phosphonates were used for P storage, P starvation would reduce the allocation of P to phosphonates. We found, however, that this strain allocates significantly more P to phosphonates relative to phosphates upon entering P-starved stationary phase (phosphonate/phosphate = 2.4) than during P-replete exponential growth (phosphonate/phosphate = 0.72; Fig. 4*B*). Cellular C-P was relatively stable across exponential and stationary phases (145 and 131, respectively); changes in the phosphonate/phosphate ratio were driven by a decrease in phosphate ester content during P starvation, while the cells kept producing phosphonates (*SI Appendix*, Fig. S10 and Table S4). Thus, P allocated to phosphonates is not internally recycled to sustain growth during periods of P starvation, consistent with our observation that *Prochlorococcus* SB lacks any known pathway to degrade and assimilate phosphonates.

Macromolecular Forms of Phosphonates. The biochemical function of phosphonates is likely linked to their macromolecular form; thus, we explored this in *Prochlorococcus* SB cultures. We expected phosphonates to be associated with polysaccharide macromolecules, because phosphonate biosynthesis genes in many *Prochlorococcus* and SAR11 genomes are surrounded with glycosylating enzymes predicted to be involved in the biosynthesis of large extracellular polysaccharide structures, such as bacterial capsules (*SI Appendix*, Fig. S5 and Table S5). Further, phosphonates are constituents of the large reservoir of dissolved polysaccharides that accumulate in the surface ocean as HMWDOP (18). However, after fractionating *Prochlorococcus* SB cells into major biochemical classes (*Materials and Methods*), we found that phosphonates were recovered not in the polysaccharide fraction but rather in the protein fraction (Fig. 4*B*)—more specifically within the methanol/acetone insoluble HMW protein fraction, which includes membrane-associated proteins. We reconcile this mismatch between expectation and findings by proposing that phosphonates are integrated into glycan polymer chains that are then posttranslationally attached to HMW membrane-anchored proteins (*SI Appendix*, Fig. S5 *A–D* and Table S5 and *SI Results*).

One of the most-common posttranslational modifications of bacterial proteins is O-linked glycosylation (45–48), in which glycan polymers are covalently bound to larger protein complexes. The phosphonate gene clusters in *Prochlorococcus* SB and two SAR11 isolate genomes are all within ± 5 Kbp of both glycan assembly enzymes and lipid carriers predicted to be involved in the biosynthesis of polysaccharide capsules (*SI Appendix*, Fig. S5 *B–D* and Table S5). In some bacteria, the glycan building blocks for capsule biosynthesis are also routed toward posttranslational O-linked glycosylation of membrane-bound lipoproteins (49). The *Prochlorococcus* SB phosphonate

gene cluster contains an initiating glycotransferase (PgIC) that links the first glycan subunit to a lipid carrier, a flippase (Wzx) that translocates the assembled glycan-bound lipid carrier to the periplasmic face, and an O-oligosaccharyltransferase (PgIL) that is the critical enzyme moving the assembled glycan chain to the final acceptor (48, 49). The PgIL family has recently been demonstrated to transfer preassembled glycan chains onto protein substrates (50, 51). These O-oligosaccharyltransferase genes are also present near *pepM* in SAR11 genomes (*SI Appendix*, Fig. S5), and 15% of GORG-Tropics SAGs contain *pepM* plus an O-oligosaccharyltransferase-like domain colocalized within the same 10-Kbp genome segment. This proportion increases to 50% if we relax the condition that the genes must occur on the same genomic contig. Thus, genomic evidence implies that phosphonates could be a common moiety involved in the posttranslational O-linked glycosylation of proteins in the ocean.

Functional Roles of Phosphonylated Glycoproteins. The widespread distribution of phosphonate synthesis genes, both taxonomically and geographically, the abundance of phosphonates in HMWDOP, and the high relative allocation of P into phosphonates in *Prochlorococcus* SB suggest that phosphonate production confers an important fitness benefit. As we argue above (*Phosphonates and P Storage*), it would appear that this benefit is not related to P nutrient storage. Rather, we hypothesize, based on several lines of inference, that this fitness advantage is linked to cell-surface defenses that mitigate mortality caused by viral lysis or protozoan grazing—two significant drivers of mortality for *Prochlorococcus* (52).

First, as detailed previously, the phosphonate biosynthesis genes are frequently flanked by genes for the biosynthesis of extracellular structures such as membrane-anchored polysaccharides. This genomic organization suggests that phosphonates are linked to polysaccharides that are posttranslationally attached to protein at the cell surface, interacting with the external environment. Many marine microbes produce extracellular layers of polysaccharides to defend against predation (53) and surface-expressed structures protect against enzymatic attack, ultraviolet radiation, phages (54–56), and protozoan grazers (57–60). Specifically, cell-surface structure modification is a common strategy to prevent phage attachment, block the injection of phage DNA (61), or influence the biochemistry of those structures. Modifications can change membrane hydrophobicity, which has been shown to influence the ingestion of *Prochlorococcus* by a nanoflagellate grazer (62), and provoke active rejection of prey or prevent digestion by inhibiting some enzymes (63). Phosphonates are often incorporated into cell-surface structures, because they protect the cell exterior from hydrolysis (64). For example, phosphonates have been found in polysaccharide B (65–68) or as a glycoprotein modification in higher organisms (69, 70). Although the bioactivity of methylphosphonate and 2-hydroxyethylphosphonate, the most-abundant phosphonates in HMWDOP, has not been investigated, many phosphonates are potent antibiotics and highly bioactive. Indeed, they can mimic phosphate and carboxylic acid functional groups, thereby interrupting important metabolic pathways (10, 71).

Second, the phosphonate biosynthesis trait in *Prochlorococcus* and SAR11 populations is present at low frequency in the wild, and its distribution is consistent with the evolutionary mechanism of negative frequency-dependent selection, which favors the presence of rare traits in a population (29). Negative frequency-dependent patterns are well documented for microbial cell-surface structures, including lipopolysaccharides, S-layers, O-antigens, and capsular polysaccharides (72). Rare variations in these surface structure traits—for example, slight structural modification—can facilitate bacterial evasion from phage and host immune responses (73, 74). However, once a variant becomes too abundant within a population, then predators, phages, or the immune system may adapt to the variant,

neutralizing any conferred advantage. Thus, natural selection in these defense traits should result in rapid gene gain and loss, which ultimately makes these genes and their variants uncommon in the wider population.

The depth distribution of phosphonate biosynthesis trait is consistent with the hypothesis that phosphonate might be used as a defense mechanism, as the relative abundance of phosphonate producers is relatively low in the surface and increases with depth, peaking at the nutrient-driven genomic transition zone/deep chlorophyll maximum layer (DCM). We hypothesize that chronic, long-term P scarcity drives the cost of phosphonate synthesis to outweigh its fitness benefit for most bacterioplankton in oligotrophic surface communities, resulting in low frequency of the phosphonate biosynthesis trait among community members. In contrast, at depths near the genomic transition zone/DCM, the greater availability of inorganic phosphate may lower the relative cost of phosphonate production while higher densities and diversity of phages (75, 76), and nanoflagellate grazers (77–79)—key grazers of *Prochlorococcus* (80) and heterotrophic bacteria (81–83)—may simultaneously increase the relative benefit of cell-surface modification with phosphonates (Fig. 5*A*).

Rare Traits Influence Major Biogeochemical Cycles. What are the consequences of phosphonate production on nutrient and other biogeochemical cycles? The depth distribution of the phosphonate synthesis trait (i.e., low abundance in the surface and higher at the DCM) mirrors P-reduction rates in the euphotic zone (84), which implies that phosphonate production is correlated to the abundance of phosphonate synthesis marker genes. This explains the virtual lack of phosphonates in surface particulate matter samples (84, 85) and implies that the amount of phosphonate in particulate organic matter should be higher at the DCM. At steady state, low rates of phosphonate production would be balanced by low rates of phosphonate degradation (C-P lyase activity) (86) and export (87). Similar reasoning argues that higher phosphonate production rates around the DCM would sustain higher overall rates of phosphonate degradation, consistent with the elevated C-P lyase gene copy numbers and C-P lyase activity measured in and below the DCM (27, 86). This implies that P redox-cycling is more active at depth than in the surface ocean. Under this scenario, open-ocean methane production by the C-P lyase degradation pathway is directly coupled to the fitness benefit of phosphonate production. This should lead to increased production of methylphosphonate, and ultimately methane, with depth, which is consistent with observations in the North Atlantic and North Pacific (88, 89). Across large areas of the ocean, this depth dependence of methylphosphonate production would in turn influence the fraction of total methane that either escapes to the atmosphere or is oxidized by methane-oxidizing bacteria (90).

If, as we hypothesize, phosphonates are used to modify cell-surface glycoproteins, the cycling of these glycoproteins might explain why phosphonates are mainly observed in HMW dissolved organic matter (i.e., linked to polysaccharides). We propose that after cell death, the labile protein component of the phosphonylated glycoproteins likely turns over quickly. However, inherently slow degradation of the structural polysaccharide component would lead to the accumulation of phosphonoglycans as HMW dissolved organic matter (91), explaining their abundance in marine DOP (Fig. 5*B*). Unlike for the majority of DOP that emerges from primary metabolism (e.g., nucleic acids, phospholipids, and energy currency) the relatively high concentrations of phosphonoglycans is a consequence of ecological interactions such as the defense mechanism we propose here.

MGI Thaumarchaeota (21), *Trichodesmium* (20), and *Prochlorococcus* (this study) isolates have all been shown to
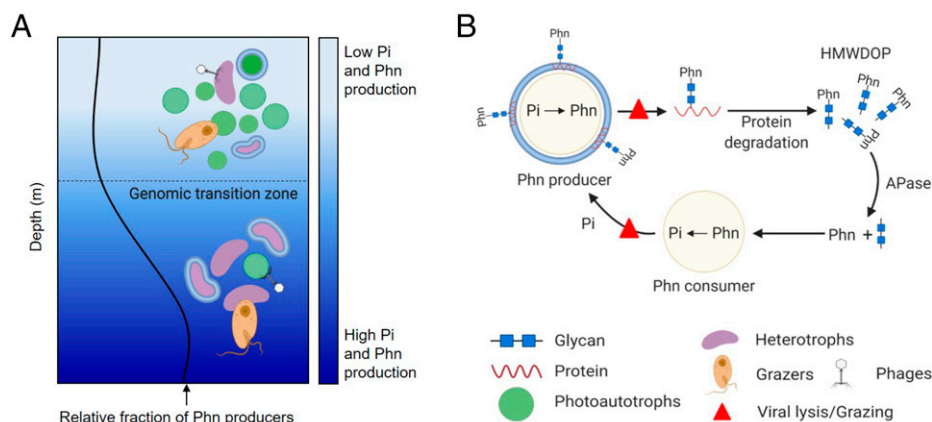
Fig. 5. Hypothesized microbial cycling of phosphonates in the surface ocean. (*A*) In the surface ocean, phosphate (Pi) is scarce and often limiting. Producing phosphonate (Phn) as a mortality defense may be costly in terms of resource allocation. Therefore, the relative fraction of phosphonate producers in the microbial community is low. As nutrient availability increases through the genomic transition zone (42), the P cost to produce phosphonate decreases and the benefit/cost ratio of phosphonate production, in which the benefit is a decrease in mortality, increases causing phosphonate producers to be relatively more abundant. (*B*) Microbes with PepM produce cell surface-layer phosphonoglycoproteins to reduce mortality from grazing or viral infection. Upon death of the cell, phosphonoglycoproteins are released into seawater where heterotrophic microbes quickly remineralize proteins leaving phosphonoglycans to accumulate as HMWDOP. Phosphonates are hydrolyzed from glycans by alkaline phosphatase (APase) family enzymes and further hydrolyzed into phosphate by C-P lyase or other hydrolytic pathways. Recycled phosphate can then be used to produce new phosphonates.

produce phosphonates. However, these groups occur at very different densities in the sunlit oligotrophic ocean. MGI Thaumarchaeota are generally low abundance ($\sim5 \times 10^4$ cells/L^{-1}) (92–94) but rapidly increase with depth in the upper mesopelagic (95). The raft morphology of *Trichodesmium* can occur at 60 colonies/m^{-3} in nonbloom conditions (96), while mean *Prochlorococcus* abundance is 4.4×10^7 cells/L^{-1} between 30°N and 30°S (97). Extrapolating from these cell densities, the fraction of each group with phosphonate production potential (this study), published P quotas from culture and field experiments (98–100), and average phosphonate content (20, 21), we estimate that MGI Thaumarchaeota, *Trichodesmium*, and *Prochlorococcus* could respectively account for <1%, <20%, and >60% of particulate phosphonate production in the global euphotic zone (*SI Appendix, SI Results*). Even though phosphonate production has not been experimentally confirmed in SAR11, it is plausible that SAR11 isolates have a cellular phosphonate content similar to *Trichodesmium* or *Prochlorococcus*. Assuming SAR11 phosphonate content is 10% of total cellular P (similar to *Trichodesmium*) and that 18% or 9.9×10^{26} cells of SAR11 in the sunlit ocean are phosphonate producers (25, 101) with P quotas of 0.16×10^{-16} mol P/$cell^{-1}$ (102), we estimate that SAR11 could easily account for 75% of the phosphonate budget in the sunlit ocean (*SI Appendix, SI Results*). Phosphonates, produced by only a minor proportion of all marine microbes, constitute a relatively large bioavailable P pool in the upper ocean. Thus, rare biosynthetic traits in numerically abundant microbes such as *Prochlorococcus* and SAR11 can have significant ecological implications for global biogeochemical cycles.

Materials and Methods

Refer to *SI Appendix* for detailed methods descriptions.

Genomic Data Sources. We used two collections of genomes for the comparative genomics analysis: MARMICRODB (103) (https://zenodo.org/record/3520509) and the GORG-Tropics dataset (24). We also used 195 surface and DCM metagenomes from Tara Oceans project (36) and 480 metagenomes from bioGEOTRACES, HOT, and BATS (35). Genome and metagenome quality control and exclusion criteria were performed as described earlier (103).

Homology Searches. Phosphonate biosynthesis and catabolism proteins were identified by homology to a collection of Hidden Markov Models using HMMER version 3.1b2 (http://hmmer.org/) and the trusted cutoffs of each

individual model. We followed an earlier approach (22) with some modifications. Briefly, we considered any genome with a valid *pepM* sequence (determined by the presence of the catalytic motif "EDK(X)5NS") as a potential phosphonate producer. We did not require the presence of *ppd*, *pdh*, or *mpnS* genes for a complete phosphonate biosynthetic pathway; we only recorded the presence of these genes if they were adjacent to *pepM*. Most *pepM*-containing genomes did have at least the decarboxylase adjacent to the *pepM* sequence or on a different contig if the cluster was on a contig break. For phosphonate catabolism proteins, gene neighborhoods were inspected to confirm the presence of multiple co-occurring genes from the PhnYZ (15), C-P lyase, 2-aminoethylphosphonate, and phosphonoacetaldehyde phosphonatase catabolism pathways. We also identified two sets of 10 highly conserved, single-copy families from *Prochlorococcus* and SAR11 lineages for use in normalization with metagenome profiling. The protein sequence profiles, results from these searches, and descriptions of the sequence families are available from https://doi.org/10.5281/zenodo.5117752. Full details are in *SI Appendix, Detailed Methods*.

Multiple Sequence Alignments, Phylogenetic Inference, and Topological Comparison. We aligned authentic PepM peptide sequences containing the "EDK (X)5NS" motif to a PepM hidden Markov model (TIGR02320), and we manually inspected the alignments to ensure the veracity of the "EDK (X)5NS" motif. We created the multi-phylum genome phylogenies from genome assemblies with confirmed phosphonate biosynthesis pathway or at least one confirmed phosphonate catabolic pathway. We used the Genome Taxonomy Database (GTDB) Toolkit version 1.3.0 pipeline (104) with default settings and the GTDB R05-RS95 database (105) to identify conserved proteins (120 bacterial proteins/122 archaeal proteins) and generate concatenated multiprotein alignments. The multi-phylum genome phylogenies and the PepM phylogeny were inferred using FastTree version 2.1.10 (106). Detailed phylogenies of *Prochlorococcus* and *Pelagibacterales*/SAR11 in *SI Appendix* were constructed as described earlier (103). We used the Mutual Clustering Information, a generalized Robinson–Foulds metric (107), for comparing the topologies of the PepM phylogeny, the genome phylogeny, and conserved gene family trees as implemented in the TreeDist version 2.0.3 R package.

Gene Enrichment Analysis. We annotated the 16 *Prochlorococcus* and 22 SAR11/*Pelagibacterales* genomes containing verified *pepM* sequences with eggNOG 4.5.1 (108) using eggNOG-Mapper version 1.0.3–3-g3e22728 (109). Significant enrichment of Kyoto Encyclopedia of Genes and Genomes pathways was determined using the hypergeometric test (110) implemented in clusterProfiler version 3.8 (111).

Identification of *pepM* Sequences within *Prochlorococcus* Genomic Islands. We predicted genomic islands using Hidden Markov Models trained from conserved gene synteny patterns in closed *Prochlorococcus* genomes as done earlier (112).

Estimation of *pepM* Prevalence in *Prochlorococcus* and SAR11 Genomes. We attempted to estimate the "true" proportion of genomes with *pepM* while correcting for genome incompleteness. Briefly, we used the estimated genome completeness from checkM version 1.0.11 (113) to estimate the number of missing bases per taxonomic group and then used this to scale the relative abundance of potential phosphonate producers per group. We estimated the corrected prevalence of phosphonate producers as:

$$\left(\frac{\sum_{i=1}^{n} p_i}{\sum_{i=1}^{n} g_i} \right) \times \sum_{i=1}^{n} \left(\frac{g_i}{c_i} \right) \times \bar{p}^{-1} \times n^{-1},$$

where for each clade, p is the length of the phosphonate biosynthesis/degradation operon in base pairs, g is the length of the genome assembly in base pairs, c is the completeness estimate from CheckM, $\bar{p}$ is the average length of all phosphonate biosynthesis/degradation operons from each clade, and n is the total number of assemblies from each clade. For the GORG datasets, we applied no additional completeness cutoffs over what was used in the original publication. Applying the commonly used completeness cutoff of 30% did not change the outcome.

Metagenome Read Classification. We identified metagenome reads derived from marker genes as reciprocal best hits using Diamond version 0.9.22.123 (114). We identified reads derived from *pepM* using empirically determined score cutoffs for the *pepM* genes present in MARMICRODB. To normalize *pepM* reads across metagenomes, we used the length normalized abundance of the 10 single-copy marker genes from *Prochlorococcus*, SAR11, and all bacteria and archaea in each metagenome.

Biotic and Abiotic Data Associated with Metagenomes. The biotic and abiotic variables used in this study were obtained and preprocessed as described in detail here: https://doi.org/10.5281/zenodo.3689249 and here: https://doi.org/10.5281/zenodo.3786232. We obtained inorganic phosphate concentrations from the GEOTRACES Intermediate Data Product IDP2017 version 3 (37), specifically from sections GA02 (115, 116), GA03, GA10 (117), and GP13. We obtained inorganic phosphate concentrations from the Tara Oceans project (118) (https://doi.pangaea.de/10.1594/PANGAEA.875579). Modeled climatological DOP and other variables were obtained from the Massachusetts Institute of Technology Darwin model (version 0.1_llc90; darwinproject.mit.edu/) and from the Simons Collaborative Marine Atlas project (119). Taxonomic profiles derived from metagenomic reads were generated as described earlier (103) using the MARMICRODB database.

Data Analysis. We fit one random forest model for *pepM* relative abundance from *Prochlorococcus*, SAR11, and bacteria plus archaea. Each model was trained using up to 44 abiotic/biotic variables including trace metal and macronutrient data from GEOTRACES and Tara Oceans, modeled climatological means from the MIT Darwin model (https://darwinproject.mit.edu/), and ecotype relative abundances. Random forest regression was implemented with the R package Ranger (120). We determined predictor variable rankings on the final model from the cross-validation step using the Boruta heuristic (121). This step allowed us to identify all predictor variables that consistently performed better than chance and to compare the importance of each variable to a reference importance level (i.e., random data). We used the R package corncob for beta-binomial regression (122) and modeled *pepM* relative abundance directly from *pepM* read counts and the median read counts to marker gene. We estimated seasonal effects in time series metagenomes using Generalized Additive Mixed Models and Linear Mixed-Effect Models implemented through the mgcv version 1.8–26 (84) nlme version 3.1–148 libraries in R v3.6.2.

Prochlorococcus Cultures under P-Replete and P-Starvation Conditions. We axenically grew *Prochlorococcus* SB and *Prochlorococcus* MIT9301 under constant light (30 μmol/quanta/m^{-2}/s^{-1}) in artificial seawater medium AMP1 prepared has described earlier (123) but using 3.75 μM [tris(hydroxymethyl)-methylamino]propanesulfonic acid (TAPS) as a buffer instead of 1 mM Hepes and a phosphate concentration of 50 μM (N/P = 16/1) and 2.28 μM (N/P = 350/1) to study the effect of P starvation. We documented the P-starvation state by adding phosphate to the culture and showing that growth resumed (Fig. 4*A*). We assessed the culture's axenicity by flow cytometry and by confirming a lack of turbidity for at least 30 d after inoculation with three purity broths: ProAC (124), MPTB (125), and ProMM (126). We cleaned all the glassware and polycarbonate bottles by soaking overnight in 2% detergent, rinsed six times with deionized water, soaked overnight in 1 M HCl, and rinsed six times with ultra–high purity water.

Cell Harvest and Treatment. We harvested the *Prochlorococcus* SB cultures grown in high N/P medium twice: ∼6 L during exponential growth and the remainder (14 L) harvested 2 d after the onset of stationary phase growth. We separated the cells from the growth medium by centrifugation (15,970 rcf for 30 min at 4 °C). After rinsing, we flash-froze the cells in liquid nitrogen and stored them at −20 °C until we conducted the NMR analyses.

NMR. We acquired the NMR spectra of whole cells at 25 °C on a 400-MHz Bruker AVANCE DPX spectrometer using a 5-mm inverse broadband probe and running TopSpin 1.3. ^{31}P shifts are reported relative to external 85% phosphoric acid at 0 ppm. For the proton-decoupled ^{31}P-NMR spectra, we used "zgdc30" with WALTZ16 decoupling and sweep width of 80 ppm, a 3-s relaxation delay, 100K scans, and 20-Hz line-broadening. We packed the *Prochlorococcus* SB and MIT9301 whole cells into a 5-mm tube (Shigemi Inc.) with magnetic susceptibility of the glass inserts matching D$_2$O. We acquired the ^{31}P-NMR spectra of *Prochlorococcus* SB protein fraction at 25 °C on a 400-MHz Bruker Ascend 400 equipped with a Sample CASE and using the program "zgpg30" with a sweep width of 80 ppm, a relaxation delay of 2 s, a 15-Hz line-broadening, and for 13K scans. ^{31}P chemical shifts are reported relative to external phosphoric acid at 0 ppm.

Elemental Composition of *Prochlorococcus* SB and MIT9301. Elemental C/N/P ratios were measured at the University of Hawai'i nutrient facility according to the protocols employed by the HOT program (https://hahana.soest.hawaii.edu/hot/methods/pcpn.html). Cell pellets from ∼900 mL of culture were transferred to combusted glass vials, dried, and powdered. C and N were measured on subsamples using a PE-2400 Carbon/Nitrogen analyzer calibrated with acetanilide standards. Cellular P was measured by the molybdenum blue method (127, 128) after combusting cell pellets at 450 °C for 3 h and dissolving the residue in 10 mL of 0.5 M HCl.

Separation of Cellular, Macromolecular Classes. To fractionate *Prochlorococcus* SB organic matter into different classes of major biochemicals, we followed the protocols of Sosa et al., 2019 (15). Cells from 3 L of culture (∼2.5 × 10^{11} cells/L) were centrifuged (12,100 rcf; 20 min) at 4 °C. The supernatant was carefully decanted, the cell pellets and concentrates combined and transferred to a 15-mL centrifuge tube, and the final suspension pelleted (12,100 rcf; 20 min) at 4 °C. The supernatant was carefully removed by pipette and the cell pellet suspended in 5 mL of cold (4 °C) 5% aqueous trichloroacetic acid (aqTCA). The suspension was agitated for 1 h, after which the mixture was centrifuged (12,100 rcf; 30 min at 4 °C) and the supernatant recovered by pipette. The aqTCA insoluble material was extracted two additional times by suspension in 5% cold (4 °C) aqTCA followed by centrifugation (12,100 rcf; 30 min at 4 °C). The aqTCA fractions were combined and evaporated to dryness by rotary evaporation to recover the polysaccharide (carbohydrate) fraction (*SI Appendix*, Fig. S13). TCA-insoluble material was suspended in 5 mL 95% aqueous ethanol (aqEtOH), manually agitated at room temperature for 20 min, and centrifuged (12,100 rcf, 30 min) and the supernatant recovered by pipette. The aqEtOH extraction was repeated two additional times, centrifuged each time (12,100 rcf; 30 min), and the supernatants recovered by pipette. The aqEtOH extracts were combined and dried by rotary evaporation (∼1 h; 35 °C) to recover the lipid fraction (*SI Appendix*, Fig. S13). Cellular material remaining in the pellet was suspended in 2.5 mL of 1 M aqNaOH and gently agitated for 1 h to hydrolyze RNA. The hydrolysis was quenched by immersing the sample tube in an ice bath for 15 min, after which the sample was acidified to pH 1 by adding 2.5 mL of 1 M aqHCl and 0.5 mL of 50% aqTCA. The mixture was allowed to rest at room temperature for 15 min to precipitate proteins and DNA. The suspension was centrifuged (12,100 rcf; 30 min at 4 °C), and the supernatant containing the RNA was removed by pipette and dried by rotary evaporation (∼1 h; 35 °C). The remaining pellet was suspended twice with 5 mL of 5% aqTCA then twice with 5 mL of 95% aqEtOH and centrifuged after each rinse (12,100 rcf; 30 min at 4 °C). The remaining solids were suspended in 5 mL of 5% aqTCA then immersed in boiling water for 30 min to hydrolyze the DNA. Immediately after, the suspension was centrifuged (12,100 rcf; 30 min at 4 °C) and the DNA-containing supernatant recovered by pipette (*SI Appendix*, Fig. S13). The residual pellet was suspended twice with 5 mL of ice-cold 5% aqTCA then twice with ice-cold 95% aqEtOH, centrifuging after each wash (12,100 rcf; 30 min at 4 °C). Finally, the pellet was hydrolyzed to recover proteins by treating the remaining cell pellet with 5 mL of 1 M aqNaOH (37 °C, 18 h). A small amount of insoluble debris remained after the protein extraction. This was removed by centrifugation (12,100 rcf; 30 min at 4 °C) and syringe filtration of the supernatant.

Protein Extraction and Precipitation. We fractionated the peptides and denatured "soluble" proteins from native HMW "insoluble" proteins using the protocol described by Hutchins et al. (129). We lysed the cell pellet from 0.5 L

right margin text
ENVIRONMENTAL SCIENCES

Acker et al.
Phosphonate production by marine microbes: Exploring new sources and potential function

PNAS | 9 of 12
https://doi.org/10.1073/pnas.2113386119

of culture (15 min, room temperature [RT]) with 1 mL of 1% sodium dodecyl sulfate (SDS) extraction buffer (1% SDS, 0.1 M Tris/HCl pH 7.5, and 10 mM ethylenediaminetetraacetic acid (EDTA) then heated at 95 °C for 10 min. We allowed the samples to cool to RT and then agitated them at 350 rpm for 1 h. We centrifuged the resulting suspension (20 min, 14,100 rcf) and decanted the supernatant. We concentrated the proteins from the supernatant by membrane centrifugation using 5K molecular weight cutoff Vivaspin units of 6 mL (Sartorius Stedim, Goettingen, Germany). We recovered the retentate (~300 µL) and added 1 mL of cold 50/50 methanol/acetone solution (acidified with HCl to a final concentration of 0.5 mM). After letting the solution rest for 3 d at −20 °C, we pelleted the insoluble proteins by centrifuge (14,100 rcf for 30 min at 4 °C) and decanted the supernatant.

Data Availability. The datasets and computer code supporting the findings in this study are available from Zenodo at https://doi.org/10.5281/zenodo.5117752.

The entire MARMICRODB dataset including a comprehensive description, peptide sequences , Kaiju version 1.6.0 (104) formatted databases, scripts, and instructions for how to use the resource is available from Zenodo at https://doi.org/10.5281/zenodo.3520509.

GEOTRACES chemical data were processed and matched to metagenome samples using code/methods available from Zenodo at https://doi.org/10.5281/zenodo.3689249.

Tara Oceans chemical and hydrographic data were processed and matched to metagenome samples using code/methods available from Zenodo at https://doi.org/10.5281/zenodo.3786232.

The list of *Prochlorococcus* core gene families and a sequence profile database are available from Zenodo at https://doi.org/10.5281/zenodo.3719132.

1. S. T. Dyhrman, "Nutrients and their acquisition: Phosphorus physiology in microalgae" in *The Physiology of Microalgae*, M. A. Borowitzka, J. Beardall, J. A. Raven, Eds. (Springer International Publishing, 2016), pp. 155–183.

2. B. A. S. Van Mooy, A. H. Devol, Assessing nutrient limitation of *Prochlorococcus* in the North Pacific subtropical gyre by using an RNA capture method. *Limnol. Oceanogr.* **53**, 78–88 (2008).

3. S. T. Dyhrman et al., The transcriptome and proteome of the diatom *Thalassiosira* pseudonana reveal a diverse phosphorus stress response. *PLoS One* **7**, e33768 (2012).

4. B. A. S. Van Mooy et al., Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature* **458**, 69–72 (2009).

5. P. Martin, S. T. Dyhrman, M. W. Lomas, N. J. Poulton, B. A. S. Van Mooy, Accumulation and enhanced cycling of polyphosphate by Sargasso Sea plankton in response to low phosphorus. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8089–8094 (2014).

6. B. Temperton, J. A. Gilbert, J. P. Quinn, J. W. McGrath, Novel analysis of oceanic surface water metagenomes suggests importance of polyphosphate metabolism in oligotrophic environments. *PLoS One* **6**, e16499 (2011).

7. E. D. Galbraith, A. C. Martiny, A simple nutrient-dependence mechanism for predicting the stoichiometry of marine ecosystems. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 8199–8204 (2015).

8. D. M. Karl, K. M. Björkman, "Dynamics of dissolved organic phosphorus" in *Biogeochemistry of Marine Dissolved Organic Matter*, D. A. Hansell, C. A. Carlson, Eds. (Academic Press, ed. 2, 2015), chap. 5, pp. 233–334.

9. L. C. Kolowith, E. D. Ingall, R. Benner, Composition and cycling of marine organic phosphorus. *Limnol. Oceanogr.* **46**, 309–320 (2001).

10. G. P. Horsman, D. L. Zechel, Phosphonate biochemistry. *Chem. Rev.* **117**, 5704–5783 (2017).

11. J. W. McGrath, J. P. Chin, J. P. Quinn, Organophosphonates revealed: New insights into the microbial metabolism of ancient molecules. *Nat. Rev. Microbiol.* **11**, 412–419 (2013).

12. B. Palenik et al., The genome of a motile marine *Synechococcus. Nature* **424**, 1037–1042 (2003).

13. A. Martinez, G. W. Tyson, E. F. Delong, Widespread known and novel phosphonate utilization pathways in marine bacteria revealed by functional screening and metagenomic analyses. *Environ. Microbiol.* **12**, 222–238 (2010).

14. S. T. Dyhrman et al., Phosphonate utilization by the globally important marine diazotroph *Trichodesmium. Nature* **439**, 68–71 (2006).

15. O. A. Sosa, J. R. Casey, D. M. Karl, Methylphosphonate oxidation in *Prochlorococcus* strain MIT9301 supports phosphate acquisition, formate excretion, and carbon assimilation into purines. *Appl. Environ. Microbiol.* **85**, e00289-19 (2019).

16. O. A. Sosa et al., Isolation and characterization of bacteria that degrade phosphonates in marine dissolved organic matter. *Front. Microbiol.* **8**, 1786 (2017).

17. P. Carini, A. E. White, E. O. Campbell, S. J. Giovannoni, Methane production by phosphate-starved SAR11 chemoheterotrophic marine bacteria. *Nat. Commun.* **5**, 4346 (2014).

18. D. J. Repeta et al., Marine methane paradox explained by bacterial degradation of dissolved organic matter. *Nat. Geosci.* **9**, 884–887 (2016).

19. R. P. Kiene, *Microbial Production and Consumption of Greenhouse Gases: Methane, Nitrogen Oxides, and Halomethanes* (American Society for Microbiology, Washington, DC, 1991), pp. 111–146.

20. S. T. Dyhrman, C. R. Benitez-Nelson, E. D. Orchard, S. T. Haley, P. J. Pellechia, A microbial source of phosphonates in oligotrophic marine systems. *Nat. Geosci.* **2**, 696–699 (2009).

21. W. W. Metcalf et al., Synthesis of methylphosphonic acid by marine microbes: A source for methane in the aerobic ocean. *Science* **337**, 1104–1107 (2012).

22. X. Yu et al., Diversity and abundance of phosphonate biosynthetic genes in nature. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20759–20764 (2013).

23. J. F. Villarreal-Chiu, J. P. Quinn, J. W. McGrath, The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment. *Front. Microbiol.* **3**, 19 (2012).

24. M. G. Pachiadaki et al., Charting the complexity of the marine microbiome through single-cell genomics. *Cell* **179**, 1623–1635.e11 (2019).

25. R. M. Morris et al., SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**, 806–810 (2002).

26. S. J. Biller, P. M. Berube, D. Lindell, S. W. Chisholm, *Prochlorococcus*: The structure and function of collective diversity. *Nat. Rev. Microbiol.* **13**, 13–27 (2015).

27. O. A. Sosa, D. J. Repeta, E. F. DeLong, M. D. Ashkezari, D. M. Karl, Phosphate-limited ocean regions select for bacterial populations enriched in the carbon-phosphorus lyase pathway for phosphonate degradation. *Environ. Microbiol.* **21**, 2402–2414 (2019).

28. M. L. Coleman, S. W. Chisholm, Ecosystem-specific selection pressures revealed through comparative population genomics. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18634–18639 (2010).

29. O. X. Cordero, M. F. Polz, Explaining microbial genomic diversity in light of evolutionary ecology. *Nat. Rev. Microbiol.* **12**, 263–273 (2014).

30. M. Ravenhall, N. Škunca, F. Lassalle, C. Dessimoz, Inferring horizontal gene transfer. *PLOS Comput. Biol.* **11**, e1004095 (2015).

31. G. C. Kettler et al., Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus. PLoS Genet.* **3**, e231 (2007).

32. M. L. Coleman et al., Genomic islands and the ecology and evolution of Prochlorococcus. *Science* **311**, 1768–1770 (2006).

33. L. S. Frost, R. Leplae, A. O. Summers, A. Toussaint, Mobile genetic elements: The agents of open source evolution. *Nat. Rev. Microbiol.* **3**, 722–732 (2005).

34. R. Feingersch et al., Potential for phosphite and phosphonate utilization by *Prochlorococcus. ISME J.* **6**, 827–834 (2012).

35. S. J. Biller et al., Marine microbial metagenomes sampled across space and time. *Sci. Data* **5**, 180176 (2018).

36. E. Karsenti et al.; Tara Oceans Consortium, A holistic approach to marine ecosystems biology. *PLoS Biol.* **9**, e1001177 (2011).

37. R. Schlitzer et al., The GEOTRACES intermediate data product 2017. *Chem. Geol.* **493**, 210–223 (2018).

38. D. M. Karl, R. Lukas, The Hawaii Ocean Time-series (HOT) program: Background, rationale and field implementation. Deep Sea Res. Part 2 Top. *Stud. Oceanogr.* **43**, 129–156 (1996).

39. D. K. Steinberg et al., Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): A decade-scale look at ocean biology and biogeochemistry. Deep Sea Res. Part 2 Top. *Stud. Oceanogr.* **48**, 1405–1447 (2001).

40. A. F. Michaels et al., Seasonal patterns of ocean biogeochemistry at the U.S. JGOFS Bermuda Atlantic time-series study site. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **41**, 1013–1038 (1994).

41. K. L. Vergin et al., High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *ISME J.* **7**, 1322–1332 (2013).

42. D. R. Mende et al., Environmental drivers of a microbial genomic transition zone in the ocean's interior. *Nat. Microbiol.* **2**, 1367–1373 (2017).

43. L. D. Quin, A. J. Williams, *Practical interpretation of P-31 NMR Spectra and Computer Assisted Structure Verification* (Advanced Chemistry Development, Inc., Toronto, ON, Canada, 2004).

44. W. W. Metcalf, W. A. van der Donk, Biosynthesis of phosphonic and phosphinic acid natural products. *Annu. Rev. Biochem.* **78**, 65–94 (2009).

45. S. C. Ku, B. L. Schulz, P. M. Power, M. P. Jennings, The pilin O-glycosylation pathway of pathogenic *Neisseria* is a general system that glycosylates AniA, an outer membrane nitrite reductase. *Biochem. Biophys. Res. Commun.* **378**, 84–89 (2009).

46. A. Vik et al., Broad spectrum O-linked protein glycosylation in the human pathogen *Neisseria gonorrhoeae. Proc. Natl. Acad. Sci. U.S.A.* **106**, 4447–4452 (2009).

47. C. M. Fletcher, M. J. Coyne, O. F. Villa, M. Chatzidaki-Livanis, L. E. Comstock, A general O-glycosylation system important to the physiology of a major human intestinal symbiont. *Cell* **137**, 321–331 (2009).

48. J. A. Iwashkiw, N. F. Vozza, R. L. Kinsella, M. F. Feldman, Pour some sugar on it: The expanding world of bacterial protein O-linked glycosylation. *Mol. Microbiol.* **89**, 14–28 (2013).

49. R. G. Lees-Miller *et al.*, A common pathway for O-linked protein-glycosylation and synthesis of capsule in *Acinetobacter baumannii*. *Mol. Microbiol.* **89**, 816–830 (2013).

50. A. Faridmoayer, M. A. Fentabil, D. C. Mills, J. S. Klassen, M. F. Feldman, Functional characterization of bacterial oligosaccharyltransferases involved in O-linked protein glycosylation. *J. Bacteriol.* **189**, 8088–8098 (2007).

51. P. M. Power, K. L. Seib, M. P. Jennings, Pilin glycosylation in *Neisseria meningitidis* occurs by a similar pathway to wzy-dependent O-antigen biosynthesis in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **347**, 904–908 (2006).

52. F. Partensky, W. R. Hess, D. Vaulot, *Prochlorococcus*, a marine photosynthetic pro-karyote of global significance. *Microbiol. Mol. Biol. Rev.* **63**, 106–127 (1999).

53. A. W. Decho, T. Gutierrez, Microbial Extracellular Polymeric Substances (EPSs) in Ocean Systems. *Front. Microbiol.* **8**, 922 (2017).

54. B. J. Wilkinson, K. M. Holmes, *Staphylococcus aureus* cell surface: Capsule as a bar-rier to bacteriophage adsorption. *Infect. Immun.* **23**, 549–552 (1979).

55. P. van der Ley, P. de Graaff, J. Tommassen, Shielding of *Escherichia coli* outer mem-brane proteins as receptors for bacteriophages and colicins by O-antigenic chains of lipopolysaccharide. *J. Bacteriol.* **168**, 449–451 (1986).

56. D. Scholl, S. Adhya, C. Merril, *Escherichia coli* K1's capsule is a barrier to bacterio-phage T7. *Appl. Environ. Microbiol.* **71**, 4872–4874 (2005).

57. C. Matz, S. Kjelleberg, Off the hook—How bacteria survive protozoan grazing. *Trends Microbiol.* **13**, 302–307 (2005).

58. M. Tarao, J. Jezbera, M. W. Hahn, Involvement of cell surface structures in size-independent grazing resistance of freshwater *Actinobacteria*. *Appl. Environ. Micro-biol.* **75**, 4720–4726 (2009).

59. E. C. Wootton *et al.*, Biochemical prey recognition by planktonic protozoa. *Environ. Microbiol.* **9**, 216–222 (2007).

60. L. Siegmund, M. Schweikert, M. S. Fischer, J. Wöstemeyer, Bacterial surface traits influence digestion by *Tetrahymena pyriformis* and alter opportunity to escape from food vacuoles. *J. Eukaryot. Microbiol.* **65**, 600–611 (2018).

61. K. D. Seed, Battling phages: How bacteria defend against viral attack. *PLoS Pathog.* **11**, e1004847 (2015).

62. B. C. Monger, M. R. Landry, S. L. Brown, Feeding selection of heterotrophic marine nanoflagellates based on the surface hydrophobicity of their picoplankton prey. *Limnol. Oceanogr.* **44**, 1917–1927 (1999).

63. K. Jürgens, C. Matz, Predation as a shaping force for the phenotypic and genotypic composition of planktonic bacteria. *Antonie van Leeuwenhoek* **81**, 413–434 (2002).

64. A. K. White, W. W. Metcalf, Microbial metabolism of reduced phosphorus com-pounds. *Annu. Rev. Microbiol.* **61**, 379–400 (2007).

65. M. J. Coyne, W. Kalka-Moll, A. O. Tzianabos, D. L. Kasper, L. E. Comstock, *Bacter-oides fragilis* NCTC9343 produces at least three distinct capsular polysaccharides: Cloning, characterization, and reassignment of polysaccharide B and C biosynthesis loci. *Infect. Immun.* **68**, 6176–6181 (2000).

66. X. Yu, N. P. J. Price, B. S. Evans, W. W. Metcalf, Purification and characterization of phosphonoglycans from *Glycomyces sp.* strain NRRL B-16210 and *Stackebrandtia nassauensis* NRRL B-16338. *J. Bacteriol.* **196**, 1768–1779 (2014).

67. A. O. Tzianabos *et al.*, The capsular polysaccharide of *Bacteroides fragilis* comprises two ionically linked polysaccharides. *J. Biol. Chem.* **267**, 18230–18235 (1992).

68. H. Baumann, A. O. Tzianabos, J. R. Brisson, D. L. Kasper, H. J. Jennings, Structural elu-cidation of two capsular polysaccharides from one strain of *Bacteroides fragilis* using high-resolution NMR spectroscopy. *Biochemistry* **31**, 4081–4089 (1992).

69. M. Urai *et al.*, Structural analysis of O-glycans of mucin from jellyfish (*Aurelia aurita*) containing 2-aminoethylphosphonate. *Carbohydr. Res.* **344**, 2182–2187 (2009).

70. B. Eckmair, C. Jin, D. Abed-Navandi, K. Paschinger, Multistep fractionation and mass spectrometry reveal zwitterionic and anionic modifications of the N- and O-glycans of a marine snail. *Mol. Cell. Proteomics* **15**, 573–597 (2016).

71. P. Kafarski, "Phosphonates: Their natural occurrence and physiological role" in *Con-temporary Topics about Phosphorus in Biology and Materials*, D. G. Churchill, M. D. Sikirić, B. Čolović, H. F. Milhofer, Eds. (IntechOpen, 2020), pp. 1–19.

72. H. Wildschutte, S. P. Preheim, Y. Hernandez, M. F. Polz, O-antigen diversity and lat-eral transfer of the wbe region among *Vibrio splendidus* isolates. *Environ. Micro-biol.* **12**, 2977–2987 (2010).

73. H. Wildschutte, D. M. Wolfe, A. Tamewitz, J. G. Lawrence, Protozoan predation, diversifying selection, and the evolution of antigenic diversity in *Salmonella*. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10644–10649 (2004).

74. B. D. Needham, M. S. Trent, Fortifying the barrier: The impact of lipid A remodelling on bacterial pathogenesis. *Nat. Rev. Microbiol.* **11**, 467–481 (2013).

75. E. Luo, J. M. Eppley, A. E. Romano, D. R. Mende, E. F. DeLong, Double-stranded DNA virioplankton dynamics and reproductive strategies in the oligotrophic open ocean water column. *ISME J.* **14**, 1304–1315 (2020).

76. E. Luo, F. O. Aylward, D. R. Mende, E. F. DeLong, Bacteriophage distributions and temporal variability in the ocean's interior. *MBio* **8**, e01903-17 (2017).

77. E. Vázquez-Domínguez *et al.*, Microbial plankton abundance and heterotrophic activity across the Central Atlantic Ocean. *Prog. Oceanogr.* **79**, 83–94 (2008).

78. E. Livanou *et al.*, Pigmented and heterotrophic nanoflagellates: Abundance and grazing on prokaryotic picoplankton in the ultra-oligotrophic Eastern Mediterra-nean Sea. *Deep Sea Res. Part 2 Top. Stud. Oceanogr.* **164**, 100 (2019).

79. U. Christaki, F. Van Wambeke, J. R. Dolan, Nanoflagellates (mixotrophs, hetero-trophs and autotrophs) in the oligotrophic eastern Mediterranean: Standing stocks, bacterivory and relationships with bacterial production. *Mar. Ecol. Prog. Ser.* **181**, 297–307 (1999).

80. L. Guillou, S. Jacquet, M. J. Chrétiennot-Dinet, D. Vaulot, Grazing impact of two small heterotrophic flagellates on *Prochlorococcus* and *Synechococcus*. *Aquat. Microb. Ecol.* **26**, 201–207 (2001).

81. T. Fenchel, Ecology of heterotrophic microflagellates. IV. Quantitative occurrence and importance as bacterial consumers. *Mar. Ecol. Prog. Ser.* **9**, 35–42 (1982).

82. J. M. Gonzalez, E. B. Sherr, B. F. Sherr, Size-selective grazing on bacteria by natural assemblages of estuarine flagellates and ciliates. *Appl. Environ. Microbiol.* **56**, 583–589 (1990).

83. J. M. Gonzàlez, Efficient size-selective bacterivory by phagotrophic nanoflagellates in aquatic ecosystems. *Mar. Biol.* **126**, 785–789 (1996).

84. B. A. S. Van Mooy *et al.*, Phosphorus cycling. Major role of planktonic phosphate reduction in the marine phosphorus redox cycle. *Science* **348**, 783–785 (2015).

85. P. Sannigrahi, E. D. Ingall, R. Benner, Nature and dynamics of phosphorus-containing components of marine dissolved and particulate organic matter. *Geo-chim. Cosmochim. Acta* **70**, 5868–5882 (2006).

86. B. N. Granzow *et al.*, A sensitive fluorescent assay for measuring carbon-phosphorus lyase activity in aquatic systems. *Limnol. Oceanogr. Methods* **19**, 235–244 (2021).

87. A. Paytan, B. J. Cade-Menun, K. McLaughlin, K. L. Faul, Selective phosphorus regen-eration of sinking marine particles: Evidence from 31P-NMR. *Mar. Chem.* **82**, 55–70 (2003).

88. O. A. Sosa *et al.*, Phosphonate cycling supports methane and ethylene supersaturation in the phosphate-depleted western North Atlantic Ocean. *Limnol. Oceanogr.* (2020).

89. S. T. Wilson, S. Ferrón, D. M. Karl, Interannual variability of methane and nitrous oxide in the north pacific subtropical gyre: Methane and nitrous oxide in the NPSG. *Geophys. Res. Lett.* **44**, 9885–9892 (2017).

90. M. A. Pack *et al.*, Methane oxidation in the eastern tropical North Pacific Ocean water column. *J. Geophys. Res. Biogeosci.* **120**, 1078–1092 (2015).

91. D. J. Repeta, "Chemical Characterization and Cycling of Dissolved Organic Matter" in *Biogeochemistry of Marine Dissolved Organic Matter*, D. A. Hansell, C. A. Carlson, Eds. (Academic Press, ed. 2, 2014).

92. A. E. Santoro *et al.*, *Thaumarchaeal* ecotype distributions across the equatorial Pacific Ocean and their potential roles in nitrification and sinking flux attenuation. *Limnol. Oceanogr.* **62**, 1984–2003 (2017).

93. A. E. Santoro *et al.*, Nitrification and nitrous oxide production in the offshore waters of the Eastern tropical south pacific. *Global Biogeochem. Cycles* **35**, e2020GB006716 (2021).

94. K. L. Casciotti, A. E. Santoro, Water column amoA and Nitrospina-like 16S rRNA gene abundances from qPCR in the Eastern Tropical South Pacific using seawater collected on R/V Atlantis (AT15-61) cruise in Jan-Feb 2010 and R/V Melville (MV1104) cruise in Mar-Apr 2011 (2021) https://doi.org/10.26008. *BCO-DMO*. **826781**, 1 (1912).

95. A. E. Santoro, R. A. Richter, C. L. Dupont, Planktonic Marine Archaea. *Annu. Rev. Mar. Sci.* **11**, 131–158 (2019).

96. C. S. Davis, D. J. McGillicuddy Jr., Transatlantic abundance of the N_2-fixing colonial cyanobacterium *Trichodesmium*. *Science* **312**, 1517–1520 (2006).

97. P. Flombaum *et al.*, Present and future global distributions of the marine Cyanobac-teria *Prochlorococcus* and *Synechococcus*. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 9824–9829 (2013).

98. T. B. Meador *et al.*, Carbon recycling efficiency and phosphate turnover by marine nitrifying archaea. *Sci. Adv.* **6**, eaba1799 (2020).

99. S. A. Sañudo-Wilhelmy *et al.*, Phosphorus limitation of nitrogen fixation by *Tricho-desmium* in the central Atlantic Ocean. *Nature* **411**, 66–69 (2001).

100. C. Grob *et al.*, Elemental composition of natural populations of key microbial groups in Atlantic waters. *Environ. Microbiol.* **15**, 3054–3064 (2013).

101. S. J. Giovannoni, SAR11 bacteria: The most abundant plankton in the oceans. *Annu. Rev. Mar. Sci.* **9**, 231–255 (2017).

102. A. E. White, S. J. Giovannoni, Y. Zhao, K. Vergin, C. A. Carlson, Elemental content and stoichiometry of SAR11 chemoheterotrophic marine bacteria. *Limnol. Ocean-ogr. Lett.* **4**, 44–51 (2019).

103. J. W. Becker, S. L. Hogle, K. Rosendo, S. W. Chisholm, Co-culture and biogeography of *Prochlorococcus* and SAR11. *ISME J.* **13**, 1506–1519 (2019).

104. P.-A. Chaumeil, A. J. Mussig, P. Hugenholtz, D. H. Parks, GTDB-Tk: A toolkit to clas-sify genomes with the Genome Taxonomy Database. *Bioinformatics*, btz848 10.1093/bioinformatics/btz848. (2019).

105. D. H. Parks *et al.*, A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).

106. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490 (2010).

107. M. R. Smith, Information theoretic generalized Robinson-Foulds metrics for com-paring phylogenetic trees. *Bioinformatics* **36**, 5007–5013 (2020).

108. J. Huerta-Cepas *et al.*, eggNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44** (D1), D286–D293 (2016).

109. J. Huerta-Cepas *et al.*, Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol. Biol. Evol.* **34**, 2115–2122 (2017).

110. I. Rivals, L. Personnaz, L. Taing, M.-C. Potier, Enrichment or depletion of a GO cate-gory within a class of genes: Which test? *Bioinformatics* **23**, 401–407 (2007).

Acker et al.
Phosphonate production by marine microbes: Exploring new sources and potential function

PNAS | 11 of 12
https://doi.org/10.1073/pnas.2113386119

111. G. Yu, L.-G. Wang, Y. Han, Q.-Y. He, clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

112. T. Hackl *et al.*, *Novel Integrative Elements and Genomic Plasticity in Ocean Ecosystems* (Cold Spring Harbor Laboratory, 2020).

113. D. H. Parks, M. Imelfort, C. T. Skennerton, P. Hugenholtz, G. W. Tyson, CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).

114. B. Buchfink, C. Xie, D. H. Huson, Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).

115. L. A. Salt, S. Heuven, M. E. Claus, E. M. Jones, H. J. W. Baar, Rapid acidification of mode and intermediate waters in the southwestern Atlantic Ocean. *Biogeosciences*, **12**, 1387–1401 (2015).

116. M. J. A. Rijkenberg *et al.*, The distribution of dissolved iron in the West Atlantic Ocean. *PLoS One* **9**, e101323 (2014).

117. N. J. Wyatt *et al.*, Biogeochemical cycling of dissolved zinc along the GEOTRACES South Atlantic transect GA10 at 40 S. *Global Biogeochem. Cycles* **28**, 44–56 (2014).

118. S. Speich *et al.*, *Environmental Context of All Samples from the Tara Oceans Expedition (2009-2013), About the Water Column Features at the Sampling Location* (PANGAEA, 2017).

119. M. D. Ashkezari *et al.*, Simons Collaborative Marine Atlas Project (Simons CMAP): An open-source portal to share, visualize, and analyze ocean data. *Limnol. Oceanogr. Methods* **19**, 488–496 (2021).

120. M. N. Wright, A. Ziegler, ranger: A fast implementation of random forests for high dimensional data in C++ and R, *J. Stat. Software* **77**, 1–17 (2017).

121. M. B. Kursa, W. R. Rudnicki, Others, Feature selection with the Boruta package. *J. Stat. Softw.* **36**, 1–13 (2010).

122. B. D. Martin, D. Witten, A. D. Willis, Modeling microbial abundances and dysbiosis with beta-binomial regression. *Ann. Appl. Stat.* **14**, 94–115 (2020).

123. L. R. Moore *et al.*, Culturing the marine cyanobacterium *Prochlorococcus: Prochlorococcus* culturing. *Limnol. Oceanogr. Methods* **5**, 353–362 (2007).

124. J. J. Morris, R. Kirkegaard, M. J. Szul, Z. I. Johnson, E. R. Zinser, Facilitation of robust growth of *Prochlorococcus* colonies and dilute liquid cultures by "helper" heterotrophic bacteria. *Appl. Environ. Microbiol.* **74**, 4530–4534 (2008).

125. M. A. Saito, J. W. Moffett, S. W. Chisholm, J. B. Waterbury, Cobalt limitation and uptake in *Prochlorococcus. Limnol. Oceanogr.* **47**, 1629–1636 (2002).

126. J. B. Waterbury, J. M. Willey, "Isolation and growth of marine planktonic cyanobacteria" in *Methods in Enzymology* (Academic Press, 1988), pp. 100–105.

127. J. D. H. Strickland, T. R. Parsons, *A Practical Handbook of Seawater Analysis* (Fisheries Research Board of Canada Bulletin 157, 2nd Edition (1972).

128. J. Murphy, J. P. Riley, A modified single solution method for the determination of phosphate in natural waters. *Anal. Chim. Acta* **27**, 31–36 (1962).

129. D. A. Hutchins *et al.*, Irreversibly increased nitrogen fixation in *Trichodesmium* experimentally adapted to elevated carbon dioxide. *Nat. Commun.* **6**, 8155 (2015).