STRUCTURAL BIOLOGY

Structural basis for methylphosphonate biosynthesis

David A. Born,[1,2]* Emily C. Ulrich,[3,4]* Kou-San Ju,[4,5,6] Spencer C. Peck,[3,4]† Wilfred A. van der Donk,[3,4,7]‡ Catherine L. Drennan[2,8,9]‡

Methylphosphonate synthase (MPnS) produces methylphosphonate, a metabolic precursor to methane in the upper ocean. Here, we determine a 2.35-angstrom resolution structure of MPnS and discover that it has an unusual 2-histidine-1-glutamine iron-coordinating triad. We further solve the structure of a related enzyme, hydroxyethylphosphonate dioxygenase from *Streptomyces albus* (*Sa*HEPD), and find that it displays the same motif. *Sa*HEPD can be converted into an MPnS by mutation of glutamine-adjacent residues, identifying the molecular requirements for methylphosphonate synthesis. Using these sequence markers, we find numerous putative MPnSs in marine microbiomes and confirm that MPnS is present in the abundant *Pelagibacter ubique*. The ubiquity of MPnS-containing microbes supports the proposal that methylphosphonate is a source of methane in the upper, aerobic ocean, where phosphorus-starved microbes catabolize methylphosphonate for its phosphorus.

Methane concentrations in oxygenated surface ocean waters are supersaturated relative to the atmosphere (*1*). This methane has been shown to derive from a biological source (*2*); however, canonical methanogenesis only occurs in obligate anaerobic archaea (*3*). Although anoxic microenvironments could provide habitats for methanogenic archaea (*4*), in situ methanogenesis has not been directly observed. Thus, the source of biological methane from the aerobic upper ocean is unknown, a conundrum referred to as the "oceanic methane paradox" (*5*). Recently, the C-P lyase pathway was identified as a biogenic source of methane in aerobic organisms prevalent in surface ocean waters (*6*, *7*). Under phosphorus stress, these organisms use the C-P lyase pathway to cleave the C-P bond in methylphosphonate (MPn), releasing methane and sequestering the phosphorus in the form of phosphate (*6*, *7*). The release of methane through metabolism of MPn has been proposed as a resolution to the oceanic methane paradox (*8*). Indeed, MPn was recently shown to be ubiquitous in dissolved organic matter from the ocean (*9*), and an MPn biosynthetic enzyme was recently identified but thus far only shown to be present in one marine microbe, the abundant archaeon *Nitrosopumilus maritimus* (*10*). The key enzyme discovered, MPn synthase (MPnS), uses molecular oxygen and a nonheme mononuclear Fe(II) center to convert 2-hydroxyethylphosphonate (2-HEP) into MPn and CO_2 (*10*) (Fig. 1). MPnS is similar in sequence to hydroxyethylphosphonate dioxygenase (HEPD) (*11*) and hydroxypropylphosphonate epoxidase (HppE) (*10*, *12*), which catalyze the formation of hydroxymethylphosphonate (HMP) from 2-HEP (Fig. 1) (*11*) and the formation of the antibiotic fosfomycin from 2-*S*-hydroxypropylphosphonate (*13*), respectively.

To establish the molecular determinants of MPnS activity and thereby establish sequence markers to use in the search for additional MPnSs, we determined x-ray structures of MPnS from *N. maritimus* in both the substrate-free and the Fe(II)- and substrate-bound states to 2.37- and 2.35-Å resolution, respectively (table S1). As expected based on sequence (fig. S1), the overall structure of MPnS is highly similar to the HEPD from *Streptomyces viridochromogenes* (*Sv*HEPD) (fig. S2) (*11*). Both enzymes are dimeric and contain the same domain structure: two β-sheet domains (β1 and β2) which together comprise a bicupin fold and two entirely α-helical domains (α1 and α2) (fig. S2A). The active site is situated in the first of the two cupin folds and completed by contributions from the N-terminal α-helical domain and the C-terminal tail of the second protomer. As noted previously (*11*), the dimeric architectures of *Sv*HEPD, and now MPnS, are analogous to the tetramer structure of HppE from *Streptomyces wedmorensis* (*12*), with each protomer of HppE corresponding to a single α-helical domain and a single β cupin fold.

Mononuclear nonheme iron enzymes are traditionally characterized by a three-residue iron-binding motif composed of two His residues and one carboxylate-bearing residue (Asp or Glu), termed the facial triad (*14*). In the active site of MPnS, a Gln residue (Q^{152}) coordinates the Fe(II) along with two His residues (H^{148} and H^{190}), forming an unusual 2-His-1-Gln facial triad (Fig. 2A and fig. S3A). The position of Gln^{152} in MPnS is distinct from that of the iron coordinating Glu (E^{176}) in *Sv*HEPD (*11*), originating from an adjacent β strand at the analogous position of the iron-coordinating Glu in HppE (*12*) (fig. S2B). The substrate 2-HEP coordinates Fe(II) in the active site of MPnS through its hydroxyl oxygen and one phosphonate oxygen (Fig. 2A and figs. S3A and S4A). The phosphonate moiety is further stabilized by interaction with Arg^{102}, Tyr^{110}, and Asn^{145} from the cupin barrel; $Lys^{28'}$ from the α-helical domain of the neighboring protomer; and $Trp^{449'}$ from the C-terminal tail of the neighboring protomer. Interestingly, the Trp^{449}-containing C-terminal tail in MPnS is mobile; in the crystal structure of substrate-free MPnS, residues Ala^{442} to Ser^{457} have an alternative position where Trp^{449} is no longer in the active site and a channel is open from the active site to bulk solvent (fig. S5, A and B). Thus, movement of the C-terminal tail and Trp^{449} appears to open and close the active site in MPnS. Although Trp^{449} is conserved between MPnS and *Sv*HEPD, the residue has not been visualized in the *Sv*HEPD structures (*11*). Three other residues of the cupin fold in MPnS (Y^{108}, I^{126}, and F^{192}) further contribute to substrate binding through hydrophobic interactions (Fig. 2A and fig. S3A).

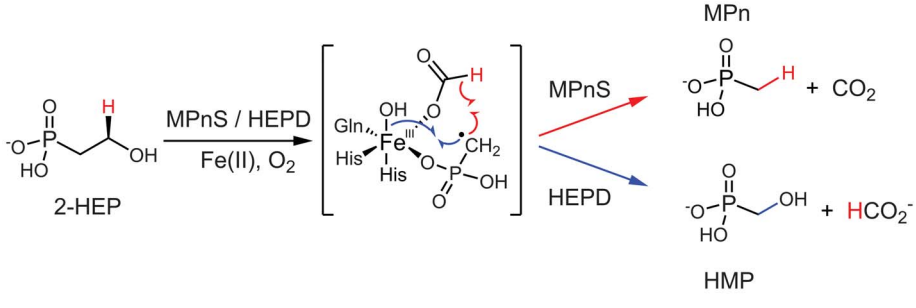

Fig. 1. The reactions of MPnS and HEPD have been proposed to proceed through a common radical intermediate. The hydrogen derived from the C2 pro-*R* position in 2-HEP is highlighted in red throughout. The radical recombination in MPnS or HEPD (*17*) is colored red or blue, respectively. For an alternative mechanism involving a cationic intermediate, see fig. S9.

[1]Graduate Program in Biophysics, Harvard University, Cambridge, MA, USA. [2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. [3]Department of Chemistry, University of Illinois at Urbana–Champaign, Urbana, IL, USA. [4]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana, IL, USA. [5]Department of Microbiology, The Ohio State University, Columbus, OH, USA. [6]Division of Medicinal Chemistry and Pharmacognosy, The Ohio State University, Columbus, OH, USA. [7]Howard Hughes Medical Institute, University of Illinois at Urbana–Champaign, Urbana, IL, USA. [8]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. [9]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA, USA.
*These authors contributed equally to this work.
†Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA.
‡Corresponding author. Email: cdrennan@mit.edu (C.L.D.); vddonk@illinois.edu (W.A.V.)

To determine whether the presence of a 2-His-1-Gln motif is predictive of MPnS activity, we selected to clone, express, and characterize a gene product from *Streptomyces albus* (strain NRRL B-16041) that displays a 2-His-1-Gln motif and 34% sequence identity to MPnS. Using a previously described phosphorus nuclear magnetic resonance (NMR) spectroscopy assay (*11*), this enzyme was shown to produce only HMP when 2-HEP is provided as substrate (fig. S6), classifying this enzyme as an HEPD. HEPDs containing a 2-His-1-Gln facial triad will hereafter be referred to as class II, whereas the previously characterized 2-His-1-Glu HEPDs will be called class I. Identification of class II HEPDs that share the same iron-ligating ligand set as MPnS demonstrates that the 2-His-1-Gln motif is not diagnostic of enzyme function, raising the question of what sequence motifs distinguish an MPnS from a class II HEPD.

To investigate the structural differences between MPnS and a class II HEPD, we determined a crystal structure of HEPD from *S. albus* (*Sa*HEPD) to 1.8-Å resolution with 2-HEP and Fe(II) bound (table S1). *Sa*HEPD has the same fold and oligomeric state as MPnS and is highly similar [root mean square deviation of 1.9 Å over 427 out of 450 C_α atoms, as calculated with PyMOL (*15*)] (fig. S2D). Consistent with the sequence alignment (fig. S1), all three iron-coordinating residues (H^{149}, Q^{153}, and H^{190} in *Sa*HEPD) and all residues forming the substrate-binding pocket (R^{103}, Y^{109}, Y^{111}, I^{127}, N^{146}, F^{192}, $K^{29'}$, and $W^{449'}$ in *Sa*HEPD) are conserved between *Sa*HEPD and MPnS (Fig. 2, A and B; fig. S3, A and B; and fig. S4B). The crystal lattice affords four independent views of this active site, and in two of the four active sites in the asymmetric unit, the electron density for Gln^{153} is consistent with an iron-coordinating conformation. However, in the other two active sites, the electron density indicates that a water molecule has replaced Gln^{153} as the ligand to Fe(II) and that Gln^{153} is now pointing away from the iron (fig. S4, D and E). This Gln is positioned such that it can hydrogen bond to Tyr^{163} that packs against Fe(II) ligand His^{190} (Fig. 2C and fig. S4E). Interestingly, Tyr^{163} is one of two residues (Gly^{184} is the other) that are close to the ligand triad and not identical between *Sa*HEPD (Y^{163}/G^{184}) and MPnS (F^{162}/I^{184}). In MPnS, Phe^{162} and Ile^{184} are in van der Waals contact with each other and fill in the area under Gln^{152} such that movement of Gln away from the Fe(II) does not appear possible (Fig. 2A). In contrast, the two residues (Y^{163} and G^{184}) in *Sa*HEPD form a cavity that can accommodate a water molecule when Gln^{153} is a ligand to Fe(II) (Fig. 2B) and can also accommodate an alternative (noncoordinating) conformation of Gln^{153} (Fig. 2C).

To investigate the importance of the alternative conformation of Gln^{153} in *Sa*HEPD, we mutated both Tyr^{163} and Gly^{184} in *Sa*HEPD to the corresponding residues in MPnS, Phe and Ile, respectively. Our structures suggest that in this double mutant the solvent cavity beneath Gln^{153} would be filled, restricting Gln^{153} to an iron-binding configuration as observed in MPnS. Indeed, the purified *Sa*HEPD-$Y^{163}F/G^{184}I$ variant catalyzes the conversion of 2-HEP to MPn with no detectable side products and with kinetic parameters [turnover number (k_{cat}), 1.2 ± 0.1 s^{-1}; and Michaelis constant for HEP ($K_{M,HEP}$), 15 ± 2 μM] very similar to those of purified wild-type *Sa*HEPD (figs. S6 and S7 and table S2). The catalytic efficiency is in fact slightly better than for *N. maritimus* MPnS (*Nm*MPnS) (k_{cat}, 0.18 ± 0.01 s^{-1}; $K_{M,HEP}$, 4.5 ± 1.1 μM) (*16*). These results demonstrate that switching the identity of only two noncoordinating active-site residues is sufficient to transform an HEPD into an MPnS. We then investigated which of these two residues is most important; *Sa*HEPD-$Y^{163}F$ produced only HMP, whereas *Sa*HEPD-$G^{184}I$ produced only MPn (fig. S8). Thus, disruption of hydrogen bonding between Tyr^{163} and Gln^{153} is insufficient to alter the reaction outcome, but preventing Gln^{153} from occupying the alternative conformation does result in the switch in product formation. We also replaced the unusual Gln iron-coordinating ligand with Glu or Ala in both *Sa*HEPD and *Nm*MPnS. *Sa*HEPD-$Q^{153}E$ was inactive, whereas the $Q^{153}A$ mutant still produced HMP, albeit with reduced efficiency (fig. S8). *Nm*MPnS-$Q^{152}A$ produced HMP with a very low level of activity, whereas *Nm*MPnS-$Q^{152}E$ was inactive under the same condition (fig. S8). Thus, removal of the Gln ligand is strongly deleterious for both MPnS and class II HEPD enzymes (for discussion, see fig. S8). Replacing the Glu ligand in class I HEPD with Gln had no effect on activity (*17*).

We next considered why restricting Gln movement would alter product formation. An iron-bound MPn-based radical and an iron-bound formate are the proposed reaction intermediates for both enzymes (Fig. 1 and fig. S9) (*18*); in MPnS, the MPn-based radical abstracts the formate hydrogen atom to yield MPn and CO_2, and in HEPD the MPn-based radical reacts with the iron-bound hydroxyl, yielding HMP and HCO_2^- (*16*, *17*). Therefore, the fate of the intermediate appears to depend directly on the geometric arrangement of formate relative to the MPn-based radical. MPnS requires the formate hydrogen atom to be in close proximity to the radical (Fig. 1). Inspection of the substrate-bound MPnS active site reveals that the positioning of an iron-bound formate intermediate would be most directly affected by the side chain of Gln^{152} (Fig. 2A and fig. S3A). With the side-chain oxygen of the Gln coordinating the iron, the side-chain amide nitrogen would be positioned appropriately to interact with the iron-bound formate, securing its position near the iron-bound MPn-radical (Fig. 3A). Thus, residues that fix the position of Gln, like a juxtaposed Ile^{184}, favor MPn production. In contrast, residues that permit Gln to adopt an alternative (non–iron-ligating) position, like a juxtaposed Gly^{184}, should support HMP production; Gln movement allows formate movement, potentially increasing the distance between the formyl hydrogen atom and the iron-bound MPn radical, disfavoring hydrogen-atom abstraction (Fig. 3B). The same arguments can be made if the reaction were to proceed via an ionic mechanism (fig. S9), except that formate would be appropriately positioned by the Gln for

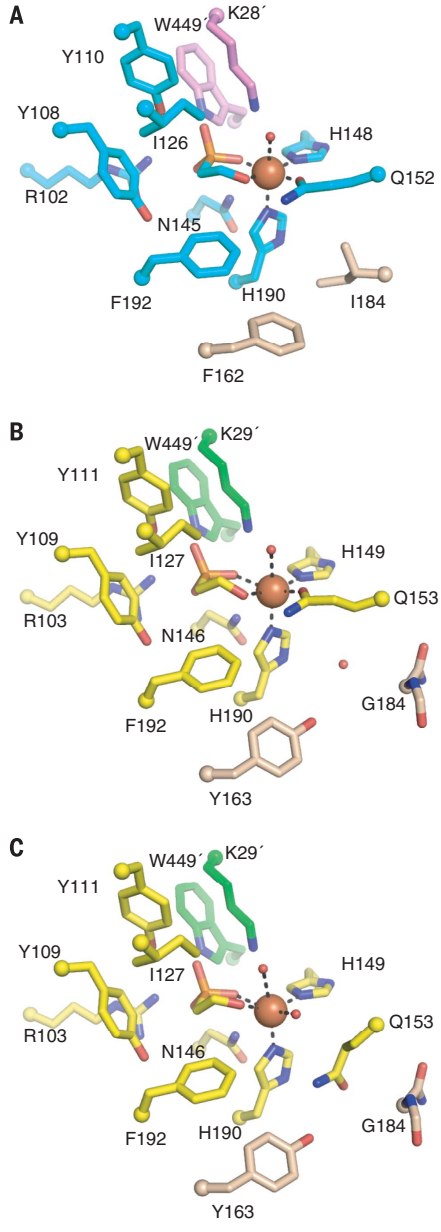

Fig. 2. Active-site structure of *Nm*MPnS and *Sa*HEPD. (**A**) The *Nm*MPnS active site is composed of residues from the β1 domain (cyan) and the α1′ domain (pink), with two key residues discussed in the text in light brown. Fe(II) (orange sphere) is ligated by substrate, Gln^{152}, His^{148}, and His^{190}. Although electron density cannot be used to distinguish oxygen from nitrogen, we are showing the oxygen of the Gln side chain coordinating Fe(II). (**B**) *Sa*HEPD active site composed of residues from the β1 domain (yellow) and the α1′ domain (green) for two protomers of the asymmetric unit in which Gln^{153} is in the iron-binding conformation. (**C**) Alternative conformation of *Sa*HEPD, where Gln^{153} is displaced from the Fe(II). Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

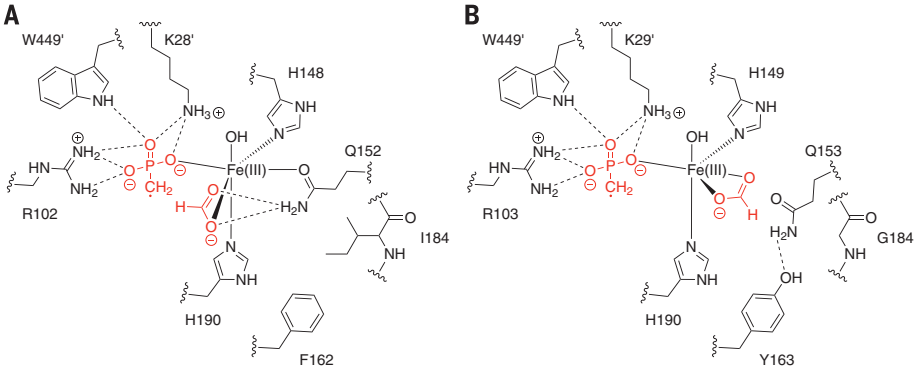

Fig. 3. Proposed intermediates for MPnS and *Sa*HEPD. (**A**) Gln^{152} in MPnS coordinates the formate intermediate, enabling H-atom abstraction by the MPn radical. (**B**) Formate displaces Gln^{153} from its iron-coordinating position in *Sa*HEPD, preventing H-atom abstraction by the MPn-based radical and promoting reaction with the iron-bound hydroxide.

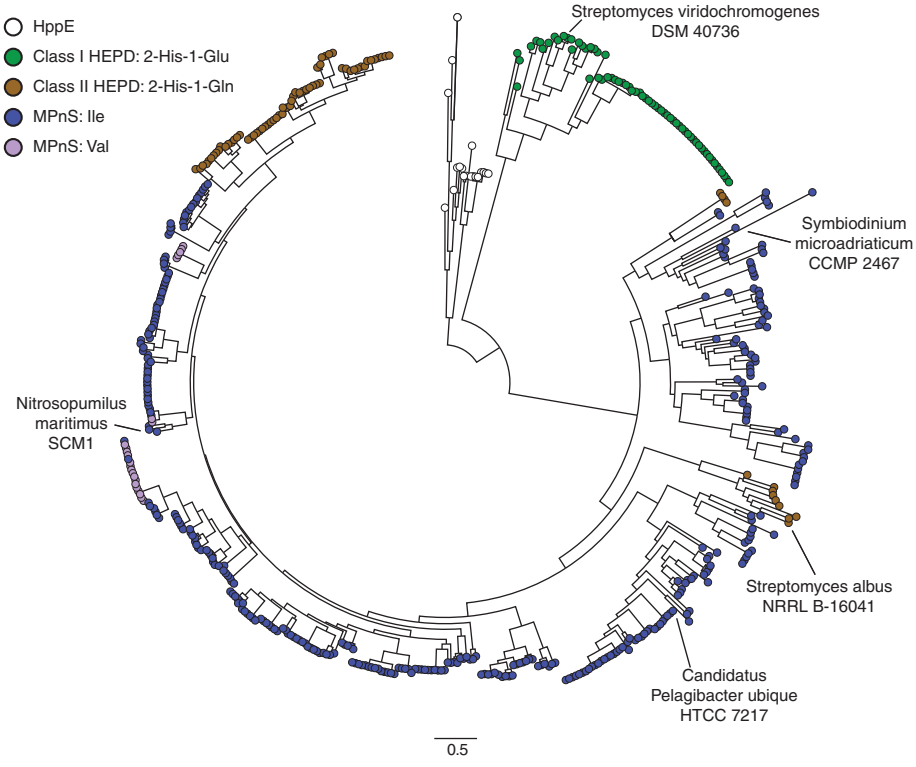

Fig. 4. Phylogeny of MPnS, HEPD, and HppE-like proteins. HEPD proteins are colored based on the iron-ligating facial triad: class I HEPD, 2-His-1-Glu (green); or class II HEPD, 2-His-1-Gln (brown). MPnS proteins are colored based on the amino acid at position 184, either Ile (blue) or Val (magenta). Complete labeling of sequence names is provided in fig. S10.

hydride rather than hydrogen atom transfer in MPnS. This model also explains the outcome of all mutant enzymes.

Next, we searched genomes for additional MPnSs. Two homologs that contain Ile at position 184 were identified in the National Center for Biotechnology Information nonredundant protein sequences database. One of these enzymes, a hypothetical protein from *Candidatus Pelagibacter ubique* strain HTCC7217 (WP_029455269.1), was expressed as a maltose-binding protein fusion enzyme in *Escherichia coli*. Indeed, this enzyme converts 2-HEP to MPn without detectable side products, consistent with its being an MPnS (fig. S6). Altogether, these results establish a sequence signature that defines MPnS enzymes: a 2-His-1-Gln facial triad with an Ile at position 184 to pack against the Gln. Phylogenetic analysis suggests that Val at position 184 may also play the same role as Ile (Fig. 4).

Extending the MPnS sequence signature to a phylogenetic analysis of genomes, metagenomes, and transcriptome shotgun assemblies reveals additional candidate MPnS-coding genes in uncultivated marine microbes and even eukaryotic marine dinoflagellates, including dinoflagellates of the genera *Symbiodinium* (Fig. 4). In contrast to HEPD-coding genes, which are present in microbial strains and metagenomes of both marine and terrestrial origin, all candidate MPnS-coding genes identified to date derive from marine environments, suggesting a relevant role for MPn in marine ecosystems. Importantly, *P. ubique* belongs to the SAR11 clade of α-proteobacteria, and the SAR11 clade, along with Thaumarchaeota (e.g., *N. maritimus*), are two of the most abundant marine microorganisms (*19*, *20*). Thus, the identification of MPnS enzymes in *P. ubique* and *N. maritimus* strongly supports a widespread role of MPn in marine carbon and phosphorus cycling, further supporting the MPn hypothesis for oceanic methane production and potentially resolving the oceanic methane paradox.

REFERENCES AND NOTES

1. M. I. Scranton, P. G. Brewer, *Deep-Sea Res.* **24**, 127–138 (1977).
2. M. E. Holmes, F. J. Sansone, T. M. Rust, B. N. Popp, *Global Biogeochem. Cycles* **14**, 1–10 (2000).
3. Y. Liu, W. B. Whitman, *Ann. N. Y. Acad. Sci.* **1125**, 171–189 (2008).
4. D. M. Karl, B. D. Tilbrook, *Nature* **368**, 732–734 (1994).
5. J. E. Rogers, W. B. Whitman, *Microbial Production and Consumption of Greenhouse Gases: Methane, Nitrogen Oxides, and Halomethanes* (American Society for Microbiology, Washington, DC, 1991).
6. C. Daughton, A. Cook, M. Alexander, *FEMS Microbiol. Lett.* **5**, 91–93 (1979).
7. P. Carini, A. E. White, E. O. Campbell, S. J. Giovannoni, *Nat. Commun.* **5**, 4346 (2014).
8. D. M. Karl *et al.*, *Nat. Geosci.* **1**, 473–478 (2008).
9. D. J. Repeta *et al.*, *Nat. Geosci.* **9**, 884–887 (2016).
10. W. W. Metcalf *et al.*, *Science* **337**, 1104–1107 (2012).
11. R. M. Cicchillo *et al.*, *Nature* **459**, 871–874 (2009).
12. L. J. Higgins, F. Yan, P. Liu, H. W. Liu, C. L. Drennan, *Nature* **437**, 838–844 (2005).
13. P. Liu *et al.*, *J. Am. Chem. Soc.* **123**, 4619–4620 (2001).
14. G. D. Straganz, B. Nidetzky, *Chem. Bio. Chem.* **7**, 1536–1548 (2006).
15. W. L. DeLano, The PyMOL Molecular Graphics System (DeLano Scientific, Palo Alto, CA, 2002).
16. H. A. Cooke, S. C. Peck, B. S. Evans, W. A. van der Donk, *J. Am. Chem. Soc.* **134**, 15660–15663 (2012).
17. S. C. Peck, J. R. Chekan, E. C. Ulrich, S. K. Nair, W. A. van der Donk, *J. Am. Chem. Soc.* **137**, 3217–3220 (2015).
18. J. T. Whitteck *et al.*, *J. Am. Chem. Soc.* **133**, 4236–4239 (2011).
19. R. M. Morris *et al.*, *Nature* **420**, 806–810 (2002).
20. M. B. Karner, E. F. DeLong, D. M. Karl, *Nature* **409**, 507–510 (2001).

ACKNOWLEDGMENTS

We thank the Giovannoni laboratory for genomic DNA from *P. ubique* HTCC7217 and H. Cooke for the plasmids for *Nm*MPnS-$Q^{152}A$ and $Q^{152}E$. This work was supported by the National Institutes of Health (GM 077596 to W.A.V.). An NIH Molecular Biophysics Training Grant T32 GM008313 supported D.A.B. C.L.D. and W.A.V. are Howard Hughes Medical Institute Investigators. ^{31}P NMR data were collected on a 600-MHz NMR spectrometer funded by NIH (S10-RR028833). This work used Northeastern Collaborative Access Team beamlines (GM103403) and a Pilatus detector (RR029205) at the Advanced Photon Source (DE-AC02-06CH11357). Structures of 2-HEP-bound Fe(II)-*Sa*HEPD, substrate-free MPnS, and 2-HEP-bound Fe(II)-MPnS are available through the Protein Data Bank under accession codes 6B9R, 6B9S, and 6B9T, respectively.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/358/6368/1336/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 and S2
References (*21*–*40*)

10 July 2017; accepted 6 November 2017
10.1126/science.aao3435

Structural basis for methylphosphonate biosynthesis

David A. Born, Emily C. Ulrich, Kou-San Ju, Spencer C. Peck, Wilfred A. van der Donk and Catherine L. Drennan

Science **358** (6368), 1336-1339.
DOI: 10.1126/science.aao3435

A source of methane in the upper ocean

Methane concentrations are high in oxygenated surface waters. Methylphosphonate (MPn) is a suggested source, but an enzyme that synthesizes Mpn (MPnS) has so far only been identified in one ocean microbe, albeit an abundant one: the archaeon *Nitrosopumilus maritimus*. Born *et al.* describe the crystal structure of MPnS and of a related enzyme that acts on the same substrate but makes a different product. By comparing the structures, they determined sequence markers that allowed them to identify MPnS in other ocean microbes, including the abundant microbe *Pelagibacter ubique* . These findings support the proposal that MPn is a source of both methane and phosphorus in the upper aerobic ocean.

Science, this issue p. 1336