



Published in final edited form as:

*Science*. 2017 December 08; 358(6368): 1336–1339. doi:10.1126/science.aao3435.

## Structural basis for methylphosphonate biosynthesis

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

<sup>1</sup>Graduate Program in Biophysics, Harvard University, Cambridge, MA, USA

<sup>2</sup>Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA

<sup>3</sup>Department of Chemistry, University of Illinois at Urbana–Champaign, Urbana, IL, USA

<sup>4</sup>Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana, IL, USA

<sup>5</sup>Department of Microbiology, The Ohio State University, Columbus, OH, USA

<sup>6</sup>Division of Medicinal Chemistry and Pharmacognosy, The Ohio State University, Columbus, OH, USA

<sup>7</sup>Howard Hughes Medical Institute, University of Illinois at Urbana–Champaign, Urbana, IL, USA

<sup>8</sup>Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA

<sup>9</sup>Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA, USA

### Abstract

Methylphosphonate synthase (MPnS) produces methylphosphonate, a metabolic precursor to methane in the upper ocean. Here we determine a 2.35-Å 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 super-saturated 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

---

\*Correspondence to: cdrennan@mit.edu and vddonk@illinois.edu.

†These authors contributed equally to this work.

‡Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA (S. C. P).

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 non-heme mononuclear Fe(II) center to convert 2-hydroxyethylphosphonate (2-HEP) into MPn and CO<sub>2</sub> (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 fosfomicin 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 non-heme 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 (Q152) coordinates the Fe(II) along with two His residues (H148 and H190), forming an unusual 2-His-1-Gln facial triad (Figs. 2A and S3A). The position of Gln152 in MPnS is distinct from that of the iron coordinating Glu (E176) 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 (Figs. 2A, S3A, and S4A). The phosphonate moiety is further stabilized by interaction with Arg102, Tyr110, and Asn145 from the cupin barrel; Lys28' from the α-helical domain of neighboring protomer; and Trp449' from the C-terminal tail of the neighboring protomer. Interestingly, the Trp449-containing C-terminal tail in MPnS is mobile; in the crystal structure of substrate-free MPnS, residues Ala442 to Ser457 have an alternative position where Trp449 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

Trp449 appears to open and close the active site in MPnS. Although Trp449 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 (Y108, I126, F192) further contribute to substrate binding through hydrophobic interactions (Fig. 2A, Fig. S3A).

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<sub>α</sub> atoms as calculated with PyMOL (15)] (Fig. S2D). Consistent with the sequence alignment (Fig. S1), all three iron-coordinating residues (H149, Q153, and H190 in *Sa*HEPD) and all residues forming the substrate-binding pocket (R103, Y109, Y111, I127, N146, F192, K29', and W449' 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 Gln153 is consistent with an iron-coordinating conformation. However, in the other two active sites, the electron density indicates that a water molecule has replaced Gln153 as the ligand to Fe(II) and that Gln153 is now pointing away from the iron (Fig. S4, D and E). This Gln is positioned such that it can hydrogen bond to Tyr163 that packs against Fe(II) ligand His190 (Fig. 2C, Fig. S4E). Interestingly, Tyr163 is one of two residues, Gly184 is the other, that are close to the ligand triad and not identical between *Sa*HEPD (Y163/G184) and MPnS (F162/I184). In MPnS, Phe162 and Ile184 are in van der Waals contact with each other and fill in the area under Gln152 such that movement of Gln away from the Fe(II) does not appear possible (Fig. 2A). In contrast, the two residues (Y163 and G184) in *Sa*HEPD form a cavity that can accommodate a water molecule when Gln153 is a ligand to Fe(II) (Fig. 2B) and can also accommodate an alternative (non-coordinating) conformation of Gln153 (Fig. 2C).

To investigate the importance of the alternative conformation of Gln153 in *Sa*HEPD, we mutated both Tyr163 and Gly184 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 Gln153 would be filled, restricting Gln153 to an iron-binding configuration as observed in MPnS. Indeed, the purified *Sa*HEPD-Y163F/G184I variant catalyzes the conversion of 2-HEP to MPn with no detectable side products and with kinetic parameters [turnover number ( $k_{\text{cat}}$ ),  $1.2 \pm 0.1 \text{ s}^{-1}$ ; and Michaelis constant for HEP ( $K_{\text{m,HEP}}$ ),  $15 \pm 2$



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  $\alpha$ -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.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

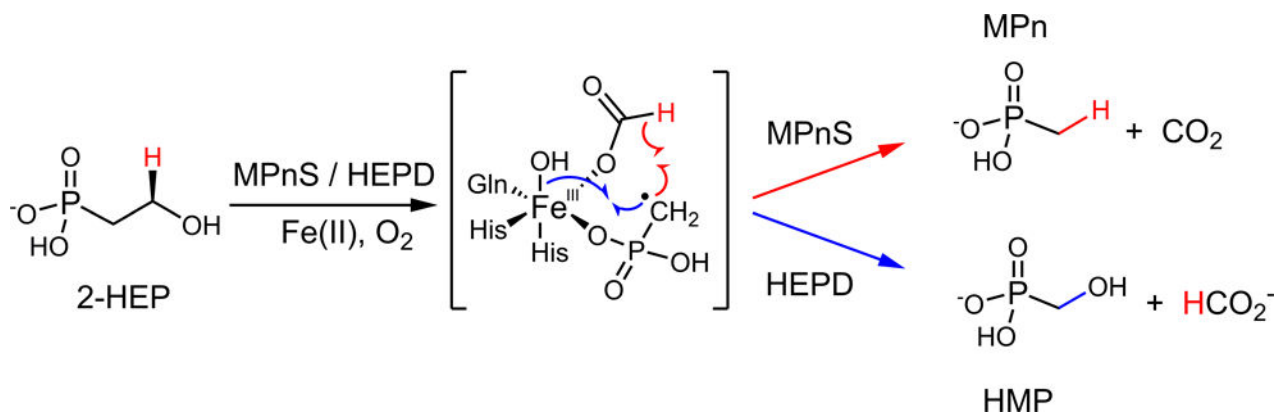
We thank the Giovannoni laboratory for genomic DNA from *P. ubique* HTCC7217 and H. Cooke for the plasmids for *Nm*MPnS-Q152A and Q152E. This work was supported by the National Institutes of Health (GM 077596 to W.A.V). An NIH Molecular Biophysics Training Grant T32 GM008313 supported D.A.B. C.L.D. and W.A.V. are Howard Hughes Medical Institute Investigators. <sup>31</sup>P NMR data were collected on a 600 MHz NMR spectrometer funded by NIH (S10-RR028833). This work used Northeastern Collaborative Access Team (NE-CAT) beamlines (GM103403) and a Pilatus detector (RR029205) at the Advanced Photon Source (DE-AC02-06CH11357). Structures of 2-HEP-bound Fe(II)-SaHEPD, 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.

## References

1. Scranton MI, Brewer PG. Occurrence of methane in the near-surface waters of the western subtropical North-Atlantic. *Deep Sea Research*. 1977; 24:127–138.
2. Holmes ME. Methane production, consumption, and air-sea exchange in the open ocean: An evaluation based on carbon isotopic ratios M. Elizabeth Holmes, 1 Francis J. Sansone, Terri M. Rust, and Brian N. Popp. *Global Biogeochemical Cycles*. 2000; 14:1–10.
3. Liu Y, Whitman WB. Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Annals of the New York Academy of Sciences*. 2008; 1125:171–189. [PubMed: 18378594]
4. Karl DM, Tilbrook BD. Production and transport of methane in oceanic particulate organic matter. *Nature*. 1994; 368:732–734.
5. Rogers JE, Whitman WB. Microbial production and consumption of greenhouse gases: methane, nitrogen oxides, and halomethanes. 1991
6. Daughton C, Cook A, Alexander M. Biodegradation of phosphonate toxicants yields methane or ethane on cleavage of the CP bond. *FEMS Microbiology Letters*. 1979; 5:91–93.
7. Carini P, White AE, Campbell EO, Giovannoni SJ. Methane production by phosphate-starved SAR11 chemoheterotrophic marine bacteria. *Nature communications*. 2014; 5
8. Karl DM, et al. Aerobic production of methane in the sea. *Nature Geoscience*. 2008; 1:473–478.
9. Repeta DJ, et al. Marine methane paradox explained by bacterial degradation of dissolved organic matter. *Nature Geoscience*. 2016; 9:884–887.

10. Metcalf WW, et al. Synthesis of methylphosphonic acid by marine microbes: a source for methane in the aerobic ocean. *Science*. 2012; 337:1104–1107. [PubMed: 22936780]
11. Cicchillo RM, et al. An unusual carbon–carbon bond cleavage reaction during phosphinothricin biosynthesis. *Nature*. 2009; 459:871–874. [PubMed: 19516340]
12. Higgins LJ, Yan F, Liu P, Liu H-w, Drennan CL. Structural insight into antibiotic fosfomycin biosynthesis by a mononuclear iron enzyme. *Nature*. 2005; 437:838–844. [PubMed: 16015285]
13. Liu P, et al. Protein purification and function assignment of the epoxidase catalyzing the formation of fosfomycin. *Journal of the American Chemical Society*. 2001; 123:4619–4620. [PubMed: 11457256]
14. Straganz G, Nidetzky B. Variations of the 2-His-1-carboxylate Theme in Mononuclear Non-Heme FeII Oxygenases. *ChemBioChem*. 2006; 7:1536–1548. [PubMed: 16858718]
15. DeLano WL. The PyMOL molecular graphics system. 2002
16. Cooke HA, Peck SC, Evans BS, van der Donk WA. Mechanistic investigation of methylphosphonate synthase, a non-heme iron-dependent oxygenase. *Journal of the American Chemical Society*. 2012; 134:15660–15663. [PubMed: 22957470]
17. Peck SC, Chekan JR, Ulrich EC, Nair SK, van der Donk WA. A common late-stage intermediate in catalysis by 2-hydroxyethyl-phosphonate dioxygenase and methylphosphonate synthase. *Journal of the American Chemical Society*. 2015; 137:3217. [PubMed: 25699631]
18. Whitteck JT, et al. On the stereochemistry of 2-hydroxyethylphosphonate dioxygenase. *Journal of the American Chemical Society*. 2011; 133:4236–4239. [PubMed: 21381767]
19. Morris RM, et al. SAR11 clade dominates ocean surface bacterioplankton communities. *Nature*. 2002; 420:806–810. [PubMed: 12490947]
20. Karner MB, DeLong EF, Karl DM. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature*. 2001; 409:507–510. [PubMed: 11206545]
21. Grant SG, Jessee J, Bloom FR, Hanahan D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proceedings of the National Academy of Sciences*. 1990; 87:4645–4649.
22. Gibson DG, et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*. 2009; 6:343–345. [PubMed: 19363495]
23. Blodgett JA, et al. Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide. *Nature chemical biology*. 2007; 3:480–485. [PubMed: 17632514]
24. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods in enzymology*. 1997; 276:307–326. [20] .
25. Schneider TR, Sheldrick GM. Substructure solution with SHELXD. *Acta Crystallographica Section D: Biological Crystallography*. 2002; 58:1772–1779. [PubMed: 12351820]
26. Abrahams J, Leslie A. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallographica Section D: Biological Crystallography*. 1996; 52:30–42. [PubMed: 15299723]
27. Vonrhein C, Blanc E, Roversi P, Bricogne G. Automated structure solution with autoSHARP. *Macromolecular Crystallography Protocols: Volume 2: Structure Determination*. 2007:215–230.
28. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*. 2004; 60:2126–2132. [PubMed: 15572765]
29. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*. 2010; 66:213–221. [PubMed: 20124702]
30. McCoy AJ, et al. Phaser crystallographic software. *Journal of applied crystallography*. 2007; 40:658–674. [PubMed: 19461840]
31. Winn MD, et al. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*. 2011; 67:235–242. [PubMed: 21460441]
32. Kabsch W. *Acta Crystallographica Section D: Biological Crystallography*. 2010; 66:125–132. XDS. [PubMed: 20124692]
33. Bunkóczi G, Read RJ. Improvement of molecular-replacement models with Sculptor. *Acta Crystallographica Section D: biological Crystallography*. 2011; 67:303–312. [PubMed: 21460448]

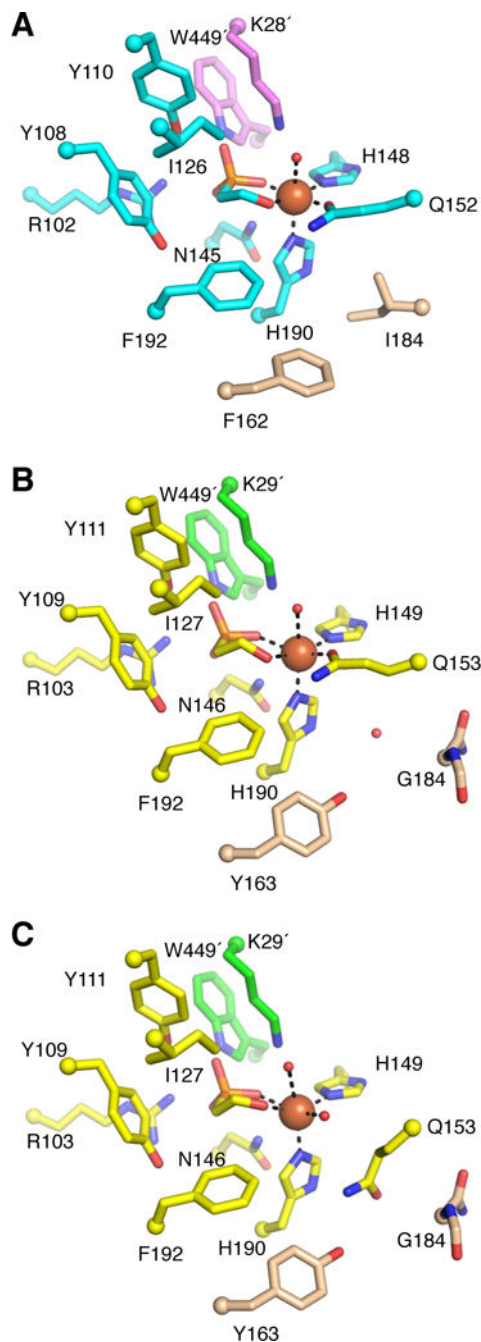
34. Morin A, et al. Collaboration gets the most out of software. *elife*. 2013; 2:e01456. [PubMed: 24040512]
35. Gasteiger E, et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research*. 2003; 31:3784–3788. [PubMed: 12824418]
36. Ju KS, et al. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proceedings of the National Academy of Sciences*. 2015; 112:12175–12180.
37. Katoh K, Misawa K, Kuma Ki, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research*. 2002; 30:3059–3066. [PubMed: 12136088]
38. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS one*. 2010; 5:e9490. [PubMed: 20224823]
39. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology*. 2007; 372:774–797. [PubMed: 17681537]
40. Peck SC, et al. O–H Activation by an Unexpected Ferryl Intermediate during Catalysis by 2-Hydroxyethylphosphonate Dioxygenase. *Journal of the American Chemical Society*. 2017; 139:2045–2052. [PubMed: 28092705]



**Figure 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 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.





**Figure 2.**

Active site structure of *NmMPnS* and *SaHEPD*. (A) The *NmMPnS* active site is composed of residues from the  $\beta 1$  domain (cyan) and the  $\alpha 1'$  domain (pink), with two key residues discussed in the text in light brown. Fe(II) (orange sphere) is ligated by substrate, Gln152, His148, and His190. 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) *SaHEPD* active site composed of residues from the  $\beta 1$  domain (yellow) and the  $\alpha 1'$  domain (green) for two protomers of the asymmetric unit in which Gln153 is in the iron-binding

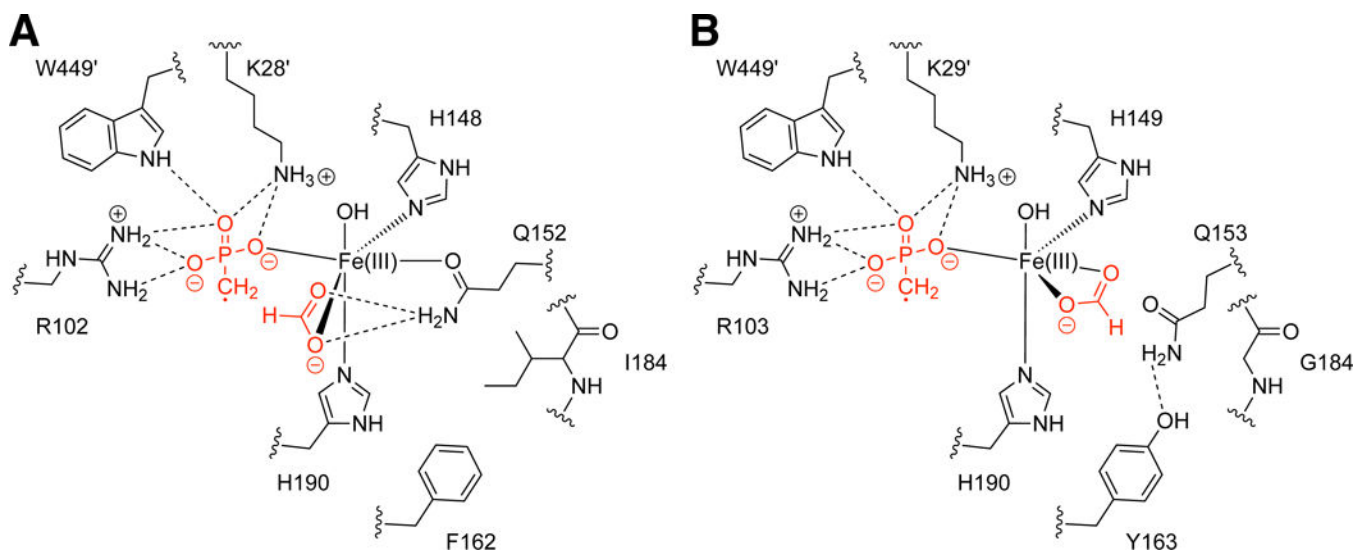
conformation. (C) Alternative conformation of *Sa*HEPD, where Gln153 is displaced from the Fe(II).

Author Manuscript

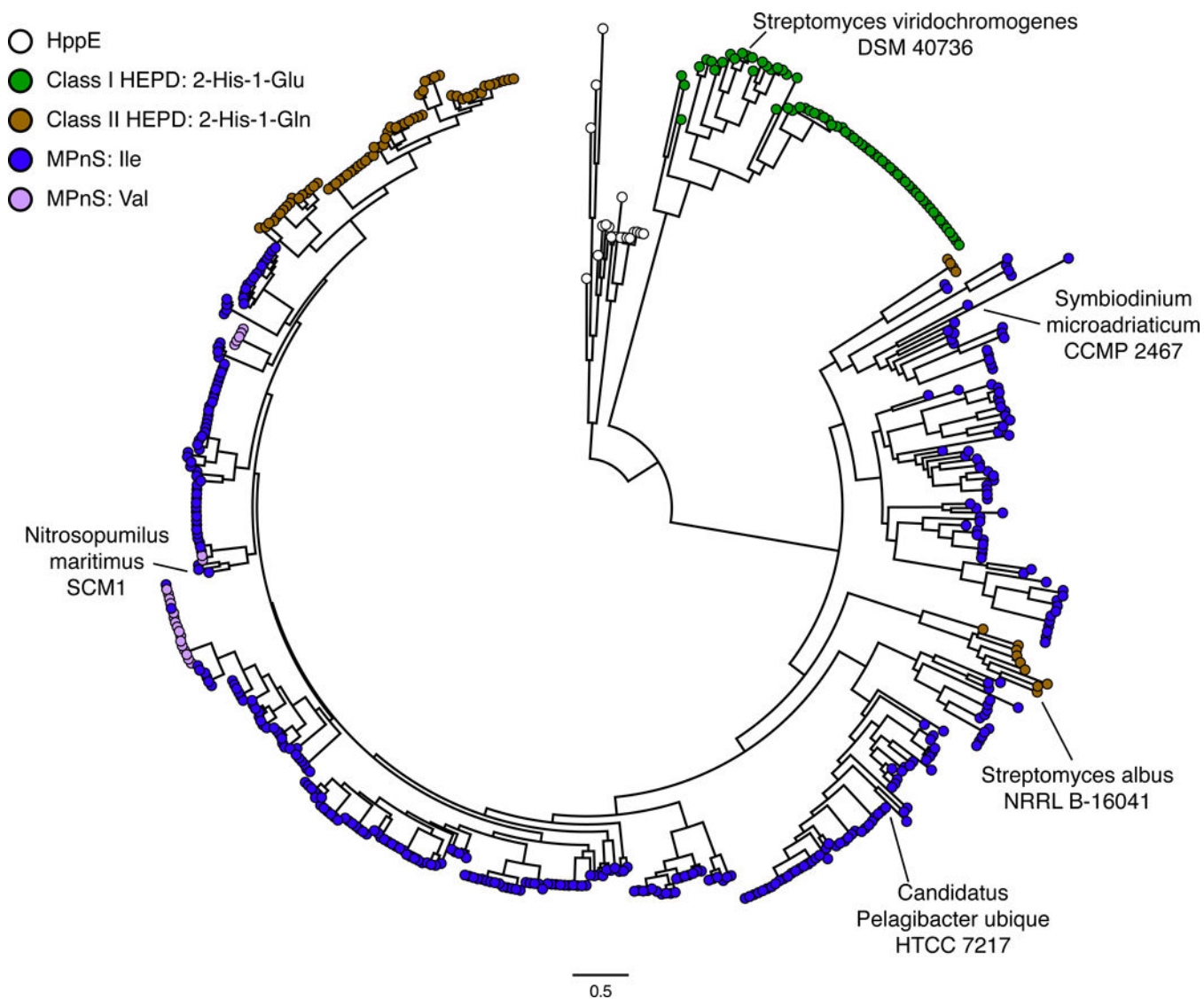
Author Manuscript

Author Manuscript

Author Manuscript



**Figure 3.** Proposed intermediates for MPnS and *Sa*HEPD. (A) Gln152 in MPnS coordinates the formate intermediate, enabling H-atom abstraction by the MPn radical. (B) Formate displaces Gln153 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.



**Figure 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.