Published in final edited form as:
Biochemistry. 2019 March 26; 58(12): 1627–1647. doi:10.1021/acs.biochem.9b00044.

A new microbial pathway for organophosphonate degradation catalyzed by two previously misannotated non-heme-iron oxygenases

Lauren J. Rajakovich[†,#], **Maria-Eirini Pandelia**[‡,⊥], **Andrew J. Mitchell**[†,‖], **Wei-chen Chang**[‡,§], **Bo Zhang**[‡,∫], **Amie K. Boal**[†,‡,*], **Carsten Krebs**[†,‡,*], and **J. Martin Bollinger Jr.**[†,‡,*]

[†]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802, United States

[‡]Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, United States

[#]Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States

[⊥]Present address: Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02453, United States

[‖]Present address: Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142

[§]Present address: Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695, United States

[∫]Present address: REG Life Sciences, LLC, South San Francisco, California 94080

Abstract

The assignment of biochemical functions to hypothetical proteins is challenged by functional diversification within many protein structural superfamilies. This diversification, which is particularly common for metalloenzymes, renders functional annotations that are founded solely on sequence and domain similarities unreliable and often erroneous. Definitive biochemical characterization to delineate functional sub-groups within these superfamilies will aid in improving bioinformatic approaches for functional annotation. We describe here the structural and functional characterization of two non-heme-iron oxygenases, TmpA and TmpB, which are encoded by a genomically clustered pair of genes found in more than 350 species of bacteria. TmpA and TmpB are functional homologues of a pair of enzymes (PhnY and PhnZ) that degrade

---

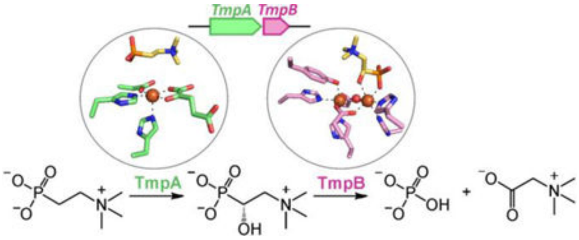
[*]**Corresponding Authors:** akb20@psu.edu, cdk10@psu.edu, jmb21@psu.edu.

ACCESSION CODES
Lc TmpA: NCBI RefSeq WP_027237574.1
PDB 6NPB, 6NPC, 6NPD
Lc TmpB: NCBI RefSeq WP_027237573.1
PDB 6NPA

SUPPORTING INFORMATION AVAILABLE
The Supporting Information includes procedures for protein purification and substrate synthesis, tables S1–2 reporting crystallographic data collection, tables S3–4, and figures S1–S25. It can be obtained online free of charge.

2-aminoethylphosphonate but instead act on its naturally occurring, quaternary ammonium analogue, 2-(trimethylammonio)ethylphosphonate (TMAEP). TmpA, an iron(II)- and 2-(oxo)glutarate-dependent oxygenase misannotated as a γ-butyrobetaine (γbb) hydroxylase, shows no activity toward γbb but efficiently hydroxylates TMAEP. The product, (*R*)-1-hydroxy-2-(trimethylammonio)ethylphosphonate [(*R*)-OH-TMAEP], then serves as the substrate for the second enzyme, TmpB. By contrast to its purported phosphohydrolytic activity, TmpB is an HD-domain oxygenase that uses a mixed-valent diiron cofactor to enact oxidative cleavage of the C–P bond of its substrate, yielding glycine betaine and phosphate. The high specificities of TmpA and TmpB for their *N*-trimethylated substrates suggests that they have evolved specifically to degrade TMAEP, which was not previously known to be subject to microbial catabolism. This study thus adds to the growing list of known pathways through which microbes break down organophosphonates to harvest phosphorus, carbon, and nitrogen in nutrient-limited niches.

Graphical Abstract



Keywords

gene annotation; enzyme; organophosphonate; iron oxygenase; HD-domain

INTRODUCTION

Members of a given metalloenzyme superfamily share a conserved protein architecture and characteristic type of metallocofactor. Even with conservation of both protein and cofactor structures, superfamily members can promote fundamentally different reaction types, transform structurally dissimilar substrates, or both.[1–4] This functional diversification limits the utility of domain conservation and sequence similarity for *de novo* prediction of metalloenzyme activity.[5–7] This problem is not unique to metalloenzymes; an estimated 98% of all gene ontology annotations are algorithmically generated and not subsequently curated,[8–10] leading to misannotations that are propagated and amplified as the number of fully sequenced genomes rapidly grows. However, given the established roles of metalloenzymes in central life processes, misannotation of these proteins, in particular, could be hindering discovery of novel functions with important biomedical or environmental implications.

Functional diversification of a single metalloprotein superfamily is exemplified by the iron(II)- and 2-(oxo)glutarate-dependent (Fe/2OG) oxygenases,[2] with ~ 100,000 members classified by the InterPro database (IPR005123, v.71.0). Fe/2OG oxygenases share a β-sandwich, cupin structural fold[11] and a largely conserved sequence motif, **H**x**D**/**E**x$_{(40–160)}$**H**, that contributes the protein ligands coordinating the characteristic mononuclear non-heme

Fe(II) cofactor.[12,13] Fe/2OG enzymes couple reduction of dioxygen to the oxidative decarboxylation of 2OG to succinate and CO_2. This reaction generates an Fe(IV)-oxo (ferryl) intermediate,[14] which, in most cases, extracts two electrons from the primary substrate to balance the complete four-electron reduction of O_2. Whereas hydroxylation is the most common outcome, halogenation, desaturation, and cyclization reactions also occur according to the same functional logic and within this single, conserved protein architecture.[15] Prediction of the reaction that a newly discovered member will catalyze is further complicated by the broad substrate scope of the superfamily, which varies from simple small molecules to complex natural products to even large macromolecules, including proteins and nucleic acids. Even high-resolution structures generally do not reveal the substrate of a given Fe/2OG oxygenase, owing to the conformational disorder of substrate-interacting loop regions often seen in the absence of the primary substrate.[11]

HD-domain proteins comprise another large and sparsely characterized superfamily of metalloenzymes,[4] with ~ 160,000 members in the InterPro database (IPR006674, v.71.0). The namesake domain of the superfamily has a characteristic $\mathbf{H}x_a\mathbf{HD}x_b\mathbf{D}$ sequence motif, which contributes the ligands for a single divalent metal ion. At the time this superfamily was first recognized, all members with known functions were hydrolases acting on phosphate esters or anhydrides,[16–19] leading to the automatic annotation of members discovered subsequently as putative phosphohydrolases.[4] However, more recent biochemical studies on several HD-domain proteins have expanded the range of known metallocofactor types and catalytic activities associated with the superfamily. Most notably, two HD-domain proteins – *myo*-inositol oxygenase (MIOX) and (*R*)-1-hydroxy-2-aminoethylphosphonate oxygenase (PhnZ) – were shown to activate dioxygen for carbon-carbon and carbon-phosphorus bond-cleavage reactions, respectively.[20–22] These enzymes harbor histidine- and carboxylate-coordinated diiron cofactors, which function in the mixed-valent Fe_2(II/III) oxidation state to activate the primary substrate and dioxygen for four-electron oxidation reactions.[20,21,23] The capacity of an HD-domain protein to bind *two* metals in the active site can be inferred from the observation of two additional, conserved histidine residues (which serve as metal ligands) in the primary structure between the aspartate residues of the canonical domain ($Hx_aHDx_b\mathbf{H}x_c\mathbf{H}x_dD$).[21,24,25] The dinuclear nature of the metallocofactor in a given HD-domain protein can thus be predicted from sequence, but this recognition is of limited value in assigning a specific activity, because a subset of proteins harboring dinuclear cofactors use them to promote the type of hydrolytic reaction for which the enzyme family was initially known.[26–28] In other words, the metal nuclearity of the cofactor does not, by itself, serve to distinguish the two major functional subclasses, phosphohydrolases and oxygenases. These recent discoveries of diversity in both cofactor structure and catalytic activity invalidate the still persistent practice of automatic annotation of newly discovered HD-domain proteins as phosphohydrolases.

The functional diversity within structural superfamilies, exemplified by these two cases, necessitates additional experimental and bioinformatic evaluation to supplement sequence- and domain-based annotations in order to improve functional predictions (of both the substrate and reaction type) for hypothetical metalloenzymes.[29–32] In bacteria, genomic context can be a valuable tool for deducing the identities of enzyme substrates and/or products.[33–36] The enzymes of a given metabolic pathway are frequently encoded by genes in a single operon or cluster,[37] allowing one to infer that the product of one enzyme will serve as the substrate for another when the corresponding genes have such synteny. A relevant example is the recently discovered pathway for degradation of the most abundant environmental organophosphonate compound, 2-aminoethylphosphonate (2-AEP). Two syntenic genes encode the enzymes PhnY and PhnZ, which enable many marine bacteria to catabolize this compound.[22] PhnY is an Fe/2OG oxygenase that hydroxylates C1 of 2-AEP to generate (*R*)-1-hydroxy-2-aminoethylphosphonate [(*R*)-OH-AEP].[22] This product serves as the substrate for PhnZ, an HD-domain diiron oxygenase, which cleaves the C–P bond to yield the final degradation products, phosphate and glycine.[21,22,24]

In this work, we identified a similar two-gene operon that encodes an Fe/2OG oxygenase and an HD-domain protein in the genomes of > 350 bacteria. In almost all cases, the Fe/2OG oxygenase, which we herein designate TmpA, had been annotated as a γ-butyrobetaine hydroxylase (BBOX). BBOXs catalyze hydroxylation of C3 of γ-butyrobetaine (γbb) to yield L-carnitine in the last step of its biosynthesis in eukaryotes[38] and in the first step of bacterial γbb degradation for carbon and nitrogen assimilation.[39,40] The HD-domain protein, herein designated TmpB, was largely annotated as a phosphohydrolase. However, the moderate sequence similarity of TmpB to the HD-domain oxygenase PhnZ (~ 30% identity), including conservation of the extended $\mathbf{H}x_a\mathbf{HD}x_b\mathbf{H}x_c\mathbf{H}x_d\mathbf{D}$ dimetal sequence motif, and its analogous genomic synteny with an Fe/2OG oxygenase led us to suspect that TmpB might actually be a new HD-domain mixed-valent diiron oxygenase (HD-MVDO). By biochemical and structural analyses, we reassigned the functions of this pair of proteins encoded in the genome of the marine bacterium *Leisingera caerulea* (*Lc*).[41,42] TmpA and TmpB catalyze the two steps of a previously unidentified pathway for degradation of the naturally occurring organophosphonate 2-(trimethylammonio)ethylphosphonate (TMAEP). TmpA hydroxylates TMAEP at C1, thereby providing the substrate for TmpB, which promotes the O_2-dependent oxidative cleavage of the C–P bond to yield glycine betaine and phosphate. Biochemical and spectroscopic evidence demonstrate that TmpA and TmpB are highly specific for their *N*-trimethylated substrates and are thus unlikely to overlap functionally with the PhnY/Z pair that degrades the corresponding unmethylated compound. The X-ray crystal structures of TmpA and TmpB rationalize this selectivity, while also confirming their similarities to BBOX and PhnZ, respectively, which was anticipated from sequence similarity and led to their initial misannotation. The biochemical and structural characterization of these enzymes, together with their genomic context, can inform functional predictions for other uncharacterized Fe/2OG and HD-domain enzymes.

EXPERIMENTAL PROCEDURES

Materials.

All chemicals used for protein over-expression and purification were purchased from Sigma-Aldrich (St. Louis, USA), unless stated otherwise.

Preparation of *Lc* TmpA, *Lc* TmpB, and *Ps* BBOX.

Detailed procedures for over-expression and purification of the proteins are provided in the Supporting Information.

Preparation of PhnY and PhnZ.

Proteins were prepared as previously described.[21]

Construction of Sequence Similarity Networks.

We generated the sequence similarity network (SSN) using the Enzyme Function Initiative – Enzyme Similarity Tool (EFI-EST) web-based server[43] and visualized it in Cytoscape (v 3.2.1).[44] We constructed a network of all sequences either (1) classified with the IPR003819 (v 59.0) “TauD/TfdA-like” domain and a length of > 350 amino acids or (2) having an *E*-value threshold $< 1 \times 10^{-10}$ in a Basic Local Alignment Search Tool (BLAST) query of the National Center for Biotechnology Information (NCBI) and the Joint Genome Institute – Integrated Microbial Genomes (JGI-IMG)[45] databases with the *Lc* TmpA sequence. Sequences with > 90% identity were binned into single nodes, and an initial pairwise alignment score threshold of $< 1 \times 10^{-50}$ was selected for vertices (Figure S1). Further network analysis was performed by increasing the threshold stringency, as described in the Supporting Information (Figures S2 and S3; Figure 2).

Synthesis of DMAEP and TMAEP.

Procedures used to synthesize 2-(dimethylamino)-ethylphosphonate (DMAEP) and 2-(trimethylammonio)ethylphosphonate (TMAEP) and characterization of the synthetic compounds are provided in the Supporting Information.

Enzymatic production of (*R*)-1-hydroxy-2-(trimethylammonio)ethylphosphonate [(*R*)-OH-TMAEP].

The (*R*)-OH-TMAEP substrate for TmpB was prepared from TMAEP by enzymatic conversion catalyzed by TmpA. A reaction (0.1 L) containing 0.005 mM TmpA, 0.01 mM $(NH_4)_2Fe(SO_4)_2$, 3 mM 2OG and 3 mM TMAEP was prepared in 25 mM ammonium acetate buffer (pH 7.5) and was incubated at room temperature with constant stirring and flushing with air. The product was purified according to the procedure used for preparation of (*R*)-OH-AEP[21] and was analyzed by NMR (Figure S4): ^{13}C NMR (D_2O) $\delta = 68.83$ (d, $J = 12.5$ Hz), 64.76 (d, $J = 145.0$), and 62.39 ppm; ^{31}P NMR (D_2O) $\delta = 13.0$ ppm; ^{1}H NMR (D_2O) $\delta = 3.60$ (m, 1H), 3.45 (m, 1H), and 3.37 (m, 1H), 1.74 (s, 9H). The stereochemistry was determined from the x-ray crystal structures of the TmpA product and TmpB substrate complexes.

Multiple-turnover enzyme assays.

End-point reactions to test for activity of Fe/2OG oxygenases with substrates and analogues contained final concentrations of 0.02 mM *apo*-TmpA (or *apo*-BBOX or *apo*-PhnY), 0.03 mM $(NH_4)_2Fe(SO_4)_2$, 0.4 mM sodium L-ascorbate, 3 mM 2OG, 2 mM substrate and 0.2 mM 2,2,3,3-[^{2}H]$_4$-succinate (d_4-succinate) in 50 mM sodium 2-[4-(2-hydroxyethyl)-piperazin-1-yl] ethanesulfonate (HEPES) buffer, pH 7.5. End-point reactions to test for activity of HD-domain proteins with substrates and analogues contained final concentrations of 0.01 mM aerobically-isolated TmpB (or aerobically-isolated PhnZ), 0.5 mM sodium L-ascorbate and 2 mM substrate in 50 mM sodium HEPES, pH 7.5. All reaction components (O_2-free) were mixed in an anoxic chamber and brought to 0.10 mL by addition of O_2-free buffer. The solutions were then removed from the chamber and the reactions were initiated by addition of 0.90 mL cold, air-saturated buffer (~ 0.4 mM O_2 at 5 °C). Reactions were incubated on ice (~3 °C), stirred, and flushed with air for 4 h. Aliquots for LC-MS analysis (0.1 mL) were treated with 0.005 mL formic acid (5% final), and aliquots for ^{31}P-NMR analysis (0.48 mL) were treated with 0.02 mL acetic acid (4% final), to terminate the reactions. Minor deviations from these general procedures are described in the appropriate figure legends.

Determination of TmpA coupling ratio.

Reaction solutions (0.1 mL) containing 0.22 mM *apo*-TmpA, 0.2 mM $(NH_4)_2Fe(SO_4)_2$, 0.4 mM TMAEP, 0.1 mM d_4-succinate and varying concentrations of 2OG (0, 0.02, 0.04, 0.06, 0.08, 0.1, 0.12, 0.14, or 0.16 mM) in 50 mM sodium HEPES, pH 7.5, were prepared in an anoxic chamber. Assay solutions were removed from the anoxic chamber and mixed with an equal volume of O_2-saturated sodium HEPES buffer, pH 7.5 (~1.8 mM O_2 at 5 °C). Reactions were quenched after 10 min by addition of 0.01 mL formic acid and were clarified by passage through centrifugal filter devices with 10 kDa molecular weight cut-off (Pall Corporation, Westborough, MA). Succinate and TMAEP were quantified by the LC-MS analysis described below.

^{31}P-NMR sample preparation and spectroscopy.

For ^{31}P-NMR measurements, sodium dithionite, ethylenediaminetetraacetic acid (EDTA), and D_2O were added to the samples to final concentrations of 0.2 mM, 0.2 mM, and 20 % (v/v), respectively. Solution ^{31}P-NMR spectra of the various phosphorus-containing compounds were recorded at room temperature on an AVX-360 or HD-500 Bruker spectrometer. The spectra were recorded with a 1D sequence with power-gated ^{1}H decoupling. Chemical shifts are quoted with respect to a phosphoric acid solution (0 ppm) prepared in the reaction buffer (sodium HEPES buffer, pH 7.5 with acetic acid, EDTA, sodium dithionite in D_2O). NMR spectra were further processed with the freely available Spinworks (version 1.3.8.1) software (Dr. Kirk Marat, University of Manitoba, Canada).

High performance liquid chromatography – mass spectrometry (LC-MS).

LC-MS analysis was carried out on an Agilent 1200 series LC system coupled to a triple quadrupole mass spectrometer (Agilent 6410 QQQ LC/MS; Agilent Technologies, Santa Clara, CA). Succinate was detected by injection of a 2-μL aliquot of a filtered sample (0.2 μm filter) onto an extend-C18 Agilent column equilibrated in 99% solvent A (water with 0.1% formic acid) and 1% solvent B (acetonitrile). Succinate was eluted from the column with an isocratic mobile phase of 99% solvent A and 1% solvent B at a flow rate of 0.4 mL/min for 10 min. Detection of succinate and the internal standard d_4-succinate was performed by electrospray ionization mass spectrometry (ESI-MS) (source parameters: gas temperature = 350 °C, gas flow = 9 L/min, nebulizer pressure = 40 psi, capillary voltage = 4000 V) in the negative mode with single ion monitoring at mass-to-charge ratios (*m/z*) of

117 and 121, respectively (fragmentor voltage = 135 V, dwell time = 200 s, delta EMV = 200 V). Succinate was quantified by multiplying the known concentration of d_4-succinate by the ratio of the integrated peak areas for succinate and d_4-succinate.

For analysis of quaternary ammonium compounds, samples (2 µL) were injected onto a SeQuant ZIC-HILIC (3.5 µm, 100 Å, PEEK 150 × 2.1 mm) column (Merck, Darmstadt, Germany) equilibrated in 2% solvent A (10 mM ammonium acetate with 0.2% formic acid) and 98% solvent B (acetonitrile). The reaction mixture was separated with a flow rate of 0.3 mL/min by applying a linear gradient of 2% - 75% solvent A over 28 min then returning to 2% solvent A over 25 min and washing with 2% solvent A for 7 min before the next injection. The compounds were detected by ESI-MS (source parameters: gas temperature = 350 °C, gas flow = 9 L/min, nebulizer pressure = 40 psi, capillary voltage = 4000 V) in the positive mode with single ion monitoring (fragmentor voltage = 135 V, dwell time = 200 s, delta EMV = 200 V) at *m/z* ratios provided in Table S1. Metabolites were quantified by comparison of integrated peak intensities to that of an internal standard (γbb or γbb-3) of known concentration.

For analysis of primary amine compounds, samples (2 µL) were injected onto a SeQuant ZIC-HILIC (3.5 µm, 100 Å, PEEK 150 × 2.1 mm) column (Merck, Darmstadt, Germany) equilibrated in 5% solvent A (water with 0.1% formic acid) and 95% solvent B (acetonitrile with 0.1% formic acid). The reaction mixture was separated with a flow rate of 0.3 mL/min by applying a linear gradient of 5% - 70% solvent A over 25 min, then returning to 5% solvent A over 25 min and washing with 5% solvent A for 10 min before the next injection. The compounds were detected by ESI-MS (source parameters: gas temperature = 350 °C, gas flow = 9 L/min, nebulizer pressure = 40 psi, capillary voltage = 4000 V) in the positive mode with single ion monitoring (fragmentor voltage = 135 V, dwell time = 200 s, delta EMV = 200 V) at *m/z* values listed in Table S2. Quantification was as described above, with ^{13}C-glycine used as the internal standard.

Stopped-flow absorption (SF-Abs) measurements and data analysis.

SF-Abs experiments were carried out at 5 °C with an Applied Photophysics Ltd. (Leatherhead, UK) SX20 stopped-flow spectrophotometer housed in an anoxic chamber (Labmaster, MBraun, Stratham, USA). The instrument was configured for single-mixing, an optical pathlength of 1 cm, and data acquisition with polychromatic light and a photodiode-array (PDA) detector. Time-dependent absorption spectra (1,000 points) were acquired with a logarithmic time base. Specific reaction conditions are provided in the appropriate figure legends.

Absorbance-versus-time traces were analyzed according to equation 1, which relates to the cycle of three consecutive, first-order, irreversible reactions shown in Scheme 1. Note that, under the conditions of the experiments (limiting O_2 and excess 2OG and TMAEP), the TmpA•Fe(II)•2OG•TMAEP reactant complex (**R**) re-forms at completion. Although the initial step (**R** → **I**) requires O_2 and is expected to be kinetically first order in this reactant, the excess of the enzyme reactant is sufficient to make the approximation of pseudo-first-order kinetics in this step an acceptably good one. Similarly, the rate of the **P** → **R** step may depend on the substrate concentrations, but both are present in large (pseudo-first-order) excess. Moreover, conversion of the product complex to the reactant complex may, as in other Fe/2OG systems, also be rate-limited by product dissociation.[46] Absorbance was monitored as a function of time at 519 nm, where **R** absorbs maximally, the TmpA•Fe(IV)=O•succin-ate•TMAEP intermediate state (**I**) is nearly isobestic with **R**, and the TmpA•Fe(II)•succinate•(*R*)-OH-TMAEP complex (**P**) is transparent, and 320 nm, where **I** absorbs intensely and **R** and **P** absorb much less intensely. In the presentation and analysis of the traces, the final absorbance was set to zero by subtraction of a constant from all absorbance values. In addition, because the first step is quite fast and nearly unresolved, k_1 was set in the regression analysis to a value of 750 s^{-1} (corresponding to a true second-order rate constant of ~ 1.5×10^6 $M^{-1}s^{-1}$). This value represents a lower limit: values greater than 750 s^{-1} gave equally good fits, but lesser values gave noticeably poorer agreement.

$$\Delta A = [\mathrm{R}]_0 \left\{ \left((\varepsilon_{\mathrm{I}} - \varepsilon_{\mathrm{R}}) \frac{k_1}{k_2 - k_1} \left(\mathrm{e}^{-k_1 t} - \mathrm{e}^{-k_2 t} \right) \right) + \right. \tag{1}$$

$$\left. \left\{ [\varepsilon_{\mathrm{P}} - \varepsilon_{\mathrm{R}}] \frac{k_2 \cdot k_1}{k_2 - k_1} \left(\frac{1}{k_3 - k_1} \left(\mathrm{e}^{-k_1 t} - \mathrm{e}^{-k_3 t} \right) - \frac{1}{k_3 - k_2} \left(\mathrm{e}^{-k_2 t} - \mathrm{e}^{-k_3 t} \right) \right) \right\} \right\}$$

Crystallization and x-ray structure solution of TmpA and TmpB.

For all structures, crystallographic datasets were collected at the GM/CA 21ID-F/G and LS-CAT 23ID-B beamlines of the Advanced Photon Source at Argonne National Lab, and the resulting diffraction images were processed with the software package HKL2000.[47] Refinement and model building were performed with the programs Refmac5[48] and COOT,[49] respectively. A summary of data collection and refinement statistics for all structures can be found in Tables S1 and S2. Ramachandran outlier analysis and other validation procedures were carried out using the Molprobity server.[50] Figures were generated in the PyMOL molecular graphics software package (Schrödinger LLC).

In an anoxic chamber at 25 °C, an O_2-free solution containing TmpA (15 mg/mL), 0.35 mM $(NH_4)_2Fe(SO_4)_2$, and 1.75 mM 2OG was mixed in a 2:1 (v:v) ratio with a precipitant solution containing of 1.0–1.4 M Li_2SO_4 and 1.6–1.8 M $(NH_4)_2SO_4$ in hanging drop vapor diffusion trials. TMAEP was soaked into existing crystals by 1:1 (v:v) mixing of crystal drops with 20 mM TMAEP prepared in the precipitant solution followed by a 4–5-h incubation at 25 °C. Crystals were prepared for data collection by cryoprotection in the precipitant solution supplemented with 19% (v/v) glycerol, mounting on rayon loops, and flash freezing in liquid nitrogen. The structure of the product complex of TmpA was obtained after adventitious exposure of crystals (described below) to O_2 during crystallization, likely due to incomplete removal of O_2 in the original screen. These crystals were obtained after mixing a solution of O_2-free TmpA (15 mg/mL), 0.35 mM $(NH_4)_2Fe(SO_4)_2$, 1.75 mM 2OG and 1.75 mM TMAEP with an equal volume of a precipitant solution (from a commercial screen by Qiagen) consisting of 1.1 M sodium malonate (pH 7.0), 0.1 M sodium HEPES (pH 7.0) and 0.5 % Jeffamine ED-2001 (pH 7.0) in a sitting-drop diffusion trial in an anoxic chamber. Crystals were prepared for data collection via soaking in the well solution supplemented with 25% (v/v) glycerol as a cryoprotectant, mounting on rayon loops, and flash freezing in liquid nitrogen.

The structure of TmpA with its phosphonate-containing product bound was solved by molecular replacement using the software package PHASER,[51] with a truncated *Hs* BBOX structure (PDB accession code 3O2G, residues 101–388) as the search model. The resulting electron density map was subjected to an autobuilding procedure using the software package ARP/wARP,[52] followed by manual model building in Coot. All other TmpA structures were solved by molecular replacement using PHASER,[51] with the structure of the TmpA product complex as the search model. The TMAEP and (*R*)-OH-TMAEP ligands were generated using the software package JLigand.[53]

The structure of the TmpA•Fe(II)•2OG complex contains residues 1–375 in chain A and residues 1–376 in chain B. The model for each chain additionally contains two residues at the N-terminus from the linker to the His_6 affinity tag. A single iron ion is modeled in chains A and B, but electron density for the 2OG co-substrate is observed only in chain A. The final model also contains 4 sulfate ions and 464 water molecules. The structure of the TmpA•Fe(II)•2OG•TMAEP complex consists of residues 1–376 in chain A, residues 1–188 and 196–376 in chain B, 2 iron ions, 1 2OG molecule, 1 TMAEP molecule (ligand TMP), 1 sulfate ion, and 521 water molecules. The model for each chain contains two or three residues at the N-terminus from the His_6 affinity tag. The TMAEP and 2OG molecules are modeled only in chain A. A loop containing residues 189–195 is disordered in chain B, but is ordered in chain A. The structure of the TmpA•product complex contains residues 1–376 in both chains A and B, 2 iron ions, 1 (*R*)-OH-TMAEP molecule (TMO), and 483 water molecules. The model for each chain contains two residues at the N-terminus from the His_6 affinity tag. Electron density for the (*R*)-OH-TMAEP product is observed solely in chain A. The electron density in the co-product site of chain A likely reflects mixed occupancy by succinate, acetate, and water molecules (Figure S15) and was therefore left unmodeled.

In an anoxic chamber, an O_2-free solution of TmpB (10 mg/mL) and 5 mM sodium L-ascorbate was mixed with an equal volume of a precipitant solution containing 0.2 M $CaCl_2$, 0.1 M sodium HEPES (pH 7.5), and 27%-33% polyethylene glycol (PEG) 4,000 in hanging drop vapor diffusion trials. (*R*)-OH-TMAEP was incorporated via 1:1 (v:v) mixing of crystal drops with 20 mM (*R*)-OH-TMAEP prepared in the crystallization solution followed by a 24-h incubation at 25 °C. Crystals were prepared for data collection by cryoprotection in the well solution supplemented with 21% (v/v) glycerol, mounting on rayon loops, and flash freezing in liquid nitrogen.

The structure of the TmpB•Fe_2•(*R*)-OH-TMAEP complex was solved by molecular replacement using the software package MOLREP[54] with the structure of PhnZ (PDB accession code 4MLM) as the search model. The final model contains four molecules of the protein in the asymmetric unit (ASU), with residues 8–195 in chain A, residues 8–194 in chain B, residues 7–76 and 85–195 in chains C and D, 6 Fe(II) ions (see explanation below), 4 Fe(III) ions, 2 (*R*)-OH-TMAEP (ligand TMO) molecules, and 261 water molecules. The diiron sites in chains A-D were modeled with one Fe(II) ion and one Fe(III) ion, because the mixed-valent cofactor accumulates to > 60% in solution when the diferric state is incubated with L-ascorbate, which was present during crystallization. However, the final oxidation state is not known and could be either Fe_2(III/III) due to adventitious oxygen exposure during crystallization or Fe_2(II/II) due to photoreduction during data collection. The (*R*)-OH-TMAEP is observed only in two of the four TmpB molecules in the ASU (chains C and D). The chains with substrate exhibit a disordered loop region, residues 77–84, near the active site. Conversely, this loop can be fully modeled in the chains that lack substrate; however, the electron density is weak and the temperature factors for this part of the model are high. The loop is found near a pair of adventitiously-bound iron ions that may be partially responsible for the ordering in the chains without substrate. This additional metal binding site, verified by an iron anomalous diffraction dataset but modeled at 50% occupancy for each site, is located between chains C and D (Figure S19).[55] Anomalous diffraction Fourier maps were generated with CAD and FFT[56–58] of the CCP4 software suite.[59]

^{57}Fe Mössbauer spectroscopy.

Mössbauer spectra were recorded on SEECO (Edina, MN) instruments that have been described previously.[14] The spectrometer used to acquire the weak-field spectra is equipped with a Janis SVT-400 variable-temperature cryostat. The external magnetic field was applied parallel to the γ beam. All isomer shifts are quoted relative to the centroid of the spectrum of α-iron foil at room temperature. Mössbauer spectra were simulated using the WMOSS spectral analysis software (www.wmoss.org, WEB Research, Edina, MN).

Continuous-Wave (CW) EPR spectroscopy.

EPR spectra at X-Band (~9.5 GHz) were acquired on a Bruker ESP-300 spectrometer equipped with an ER/4102 ST resonator (Bruker), an Oxford Instruments continuous-helium-flow cryostat, and an Oxford Instruments temperature controller (ITC 502). For all experiments, quartz tubes with 3 mm inner and 4 mm outer diameters were used (QSI).

RESULTS

The Fe/2OG oxygenase TmpA is a TMAEP hydroxylase

Discovery of the TMAEP hydroxylation activity of TmpA.—The *Lc* TmpA protein shares 26% sequence identity with *Homo sapiens* (*Hs*) BBOX and 33% identity with *Pseudomonas sp. AK-1* (*Ps*) BBOX. The latter enzyme is, to the best of our knowledge, the only characterized bacterial BBOX orthologue.[39,60–62] To evaluate the annotation of TmpA as a BBOX, the enzyme was tested in a multiple-turnover reaction for the ability to hydroxylate γbb. TmpA was found neither to consume γbb nor to produce the hydroxylated product, L-carnitine (Figure 1). By contrast, the positive-control enzyme, *Ps* BBOX, completely converted γbb to L-carnitine under the same reaction conditions (Figure 1).

The verified γbb hydroxylation activity of *Ps* BBOX and complete absence of any such activity for TmpA prompted further scrutiny and comparison of their primary structures. A sequence similarity network (SSN)[43,63] was constructed for a Fe/2OG superfamily subset that included BBOX sequences and TmpA-like sequences (Figure S1). This subset also included the biochemically characterized enzyme, ε-trimethyllysine hydroxylase (TMLH), which catalyzes the first step in eukaryotic L-carnitine biosynthesis.[38,64] TMLH, a close homologue of BBOX, was used as a rational benchmark for selection of a minimum pairwise alignment score that generates isofunctional clusters (acting on the same substrate). Separation of the isofunctional TMLH cluster from the BBOX cluster was achieved at a pairwise alignment score of 10^{-65} (Figure S2). At this threshold, the BBOX and TmpA-like sequences also segregate into distinct sub-clusters, with the nodes corresponding to the *Ps* BBOX and *Lc* TmpA proteins no longer connected via an edge vector because their pairwise alignment score exceeds 10^{-65}. The separation of *Ps* BBOX and *Lc* TmpA into presumptively isofunctional clusters is consistent with the experimental observation that they do not act on the same substrate. In a network constructed from only bacterial sequences and with a slightly higher stringency for edges (Figure S3), the clusters containing nodes for *Ps* BBOX and *Lc* TmpA become more isolated and adopt a clustering pattern that reflects genomic context (Figure 2). Yellow nodes, which generally cluster with *Lc* TmpA, represent proteins that are encoded adjacent to a gene encoding a TmpB-like HD-domain protein and often downstream of a gene encoding a LysR-type transcription regulator (Figure 2, yellow box). Within this cluster, a number of sequences are apparent protein fusions of TmpA- and TmpB-like domains (orange nodes). Conversely, genes encoding proteins represented by gray and green nodes lack the neighboring gene encoding an HD-domain protein. Proteins represented by the green nodes are encoded in gene neighborhoods similar to the γbb-degradation operon previously characterized in *Ps*.[39] This operon encodes the *Ps* BBOX enzyme, as well as proteins involved in betaine uptake and further degradation of L-carnitine to glycine betaine (Figure 2, green box).[39] Thus, the proteins represented by green nodes are likely to be authentic BBOX enzymes that enable bacterial γbb assimilation. The clear dichotomy between the genomic contexts of TmpA-clustered proteins and those clustered with *Ps* BBOX is consistent with their use of different substrates, as shown experimentally for *Lc* TmpA.

We initially predicted that the substrate for TmpA might be structurally similar to the γbb substrate of BBOX and therefore tested various analogues of that compound. None of the analogues most similar in structure to γbb, possessing both the trimethylammonium and carboxylate groups, were consumed by TmpA (Figure S5). We noted that the BBOX active site residues that interact with the quaternary amine of γbb (*vide infra*) are, in fact, conserved (or conservatively substituted) in TmpA sequences. We therefore considered that the TmpA substrate would retain this structural element but have a different substituent replacing the carboxylate group. TmpA was found to act, in the presence of its presumptive co-substrates, 2OG and O_2, on the quaternary ammonium-containing compound phosphocholine (Figure S6). The analogy of the genomic synteny of the TmpA/B pair to that of the phosphonate-degrading PhnY/Z pair suggested the possibility that the former pair might also degrade a phosphonate compound. Thus, we tested the naturally occurring phosphonate analogue of phosphocholine, 2-(trimethylammonio)ethylphosphonate (TMAEP) and observed efficient consumption (200 turnovers) under the same assay conditions (Figure 3). TMAEP was completely converted to a product with a distinct ^{31}P-NMR chemical shift and an *m/z* of +16 relative to that of the substrate, suggestive of substrate hydroxylation (Figure 3). By NMR analysis (Figure S4), we determined the purified product to be 1-hydroxy-2-(trimethylammonio)ethylphosphonate (OH-TMAEP).

The configuration of C1 (*R*) was evident from the X-ray crystal structures of the TmpA•product and TmpB•substrate complexes (described below). The C1 configuration matches that generated by PhnY in the cognate intermediate of the previously identified pathway for degradation of the unmethylated analog, 2-AEP.[22] Hydroxylation of TMAEP was dependent on the presence of TmpA, 2OG, and O_2 (Figure S7), and tight coupling between succinate production and TMAEP consumption was observed, with a stoichiometry close to the theoretical value of unity (Figure S7). By contrast, no conversion of TMAEP under analogous conditions was observed upon incubation with PhnY (Figure 3), which natively hydroxylates the corresponding primary amine analogue (2-AEP). *Ps* BBOX converted only a small fraction (< 10%) of the TMAEP in a 4-h incubation (Figure 3).

As a foundation for assessing the enzyme's specificity for this substrate, the capacity of TMAEP to bind to TmpA and trigger its reaction with O_2 were probed by UV-visible absorption spectroscopy. In the absence of substrate, the ternary TmpA•Fe(II)•2OG complex exhibited the usual visible absorption feature at 530 nm (Figure 4A), attributed in previous studies on other Fe/2OG oxygenases to a metal-to-2OG charge-transfer transition.[65] Mixing of this complex with limiting O_2 resulted in only a very sluggish reaction, marked by slow development of an absorption feature centered at ~ 350 nm (Figure S8). These spectral changes likely reflect unproductive oxidation of the Fe(II) cofactor to Fe(III), as observed for other Fe/2OG enzymes.[14,66] Addition of TMAEP to an anoxic solution of the ternary TmpA•Fe(II)•2OG complex elicited additional structure in the MLCT band, as well as small but reproducible hyper- and hypsochromic shifts (Figure 4A), reflecting the effect of TMAEP binding on the geometry and electronic structure of the cofactor.[67–69] Mixing of the resultant quaternary TmpA•Fe(II)•2OG•TMAEP complex with limiting O_2 led to rapid, transient changes reflective of a single enzyme turnover (Figure 4B). The 520-nm feature of the reactant complex decayed and re-developed (Figure 4B, C), signifying conversion of the reactant complex to a transparent state – most likely the TmpA•Fe(II)•succinate•(*R*)-OH-TMAEP product complex[70] – followed by re-formation of the reactant complex. Non-linear regression analysis of the kinetic traces (described in the Experimental Section) yielded apparent first-order rate constants of 59 s^{-1} for the decay phase and 2.0 s^{-1} for the re-development phase. Additionally, a feature with an apparent maximum of 318 nm developed by the time of the first reliable spectrum (4 ms; Figure 4B, red trace) and decayed with a rate constant (67 s^{-1}) not significantly different from the larger value extracted from analysis of the 520-nm traces (Figure 4C). A final re-development phase had an observed rate constant equivalent to that for re-development of the visible feature of the reactant complex. These transient UV absorption spectra and kinetic traces reflect the rapid (nearly in the dead-time of the stopped-flow instrument) accumulation of the ferryl intermediate, its slower decay to a product complex, and the overall rate-limiting re-formation of the reactant complex upon dissociation of the products and re-binding of 2OG and TMAEP (Scheme 1). We have previously documented similar kinetic schemes for multiple other Fe/2OG oxygenases.[14,66,71–74]

Examining the substrate specificity of TmpA.—Structural analogues of TMAEP were tested for their ability to bind to TmpA, trigger O_2 activation, and serve as substrates in multiple-turnover activity assays. A panel of natural and synthetic organophosphonate compounds was screened to examine whether the phosphonate alone was sufficient for substrate processing. TmpA was unable to consume the vast majority of these compounds, with the exceptions of 2-AEP and 3-aminopropylphosphonate (the latter not known to occur in nature), which were processed only very inefficiently (Figures S9, S10). Compounds with a hydroxyl, carboxylate, or sulfonate group in place of the phosphonate (e.g., choline, γbb, γbb-3, taurine, γ-aminobutyric acid, and β-alanine) were also not detectably consumed after a 4-h incubation with TmpA, Fe(II), and co-substrates (Figure S10). Addition of any one of the latter set of analogues to the TmpA•Fe(II)•2OG ternary complex did not perturb the absorption spectrum (Figure S11), suggesting that none binds with affinity comparable to that of TMAEP. Furthermore, none of these analogues triggered the reaction with O_2, demonstrated by their failure to promote either the rapid loss of the 520-nm feature of the reactant complex or the development of the transient 318-nm absorption feature of the ferryl intermediate (Figure S11).

Analogues retaining the phosphonate moiety but having reduced degrees of amine methylation were also examined as substrates for TmpA and were found to be poorer than TMAEP. Addition of 2-(dimethylamino)ethylphosphonate (DMAEP) to the TmpA•Fe(II)•2OG ternary complex induced a shift in the absorption spectrum similar to that caused by TMAEP (Figure 5A, blue). Subsequent mixing of the TmpA•Fe(II)•2OG•DMAEP complex with limiting O_2 resulted in the spectral changes characteristic of a single turnover, as observed for TMAEP (Figure 5B, blue). However, the observed rate constant for the decay phase at 520 nm was approximately four times less (k_{obs} = 13 s^{-1}) than that observed with TMAEP. A new species with an increased *m/z* of +16 relative to that of the substrate was detected by LC-MS (Figure S10), and the presence of a new peak with a distinct ^{31}P-NMR chemical shift (Figure 5C) suggested hydroxylation of the dimethyl analogue. In multiple-turnover reactions with low concentrations of TmpA, the turnover frequency with DMAEP (v_0/[E] = 0.43 s^{-1}) was 1.6-fold less than with TMAEP (v_0 = 0.67 s^{-1}). Such a modest difference in rate compelled us to assess more directly the enzyme's preference for the *N*-trimethyl versus the *N*-dimethyl substrate in a competition assay with the two compounds present in equal concentration. TmpA consumed exclusively TMAEP at early incubation times and began to process DMAEP only after TMAEP was almost entirely depleted (Figure 6B), consistent with a preference for the *N*-trimethyl substrate. With the corresponding unmethylated compound, 2-AEP, TmpA was found to carry out fewer than 10 turnovers after a 4-h incubation with excess 2OG and O_2 (Figure 5C and S11). Accordingly, addition of 2-AEP to the TmpA•Fe(II)•2OG complex elicited a less pronounced shift in its absorption spectrum than for TMAEP or DMAEP (Figure 5A, green), and no detectable triggering of the O_2 reaction was observed (Figure 5B, green). These comparisons show that the presence of three *N*-methyl groups is important for high-affinity, functional binding in the active site.

As noted previously, TmpA degrades the phosphoester analogue of TMAEP, phosphocholine (PC). The reaction produces phosphate and glycine betaine aldehyde, the former determined by its ^{31}P-NMR chemical shift and the latter by its *m/z* in LC-MS analysis (Figure S6). The observed products presumably result from hydroxylation at C1 of PC followed by elimination of phosphate. This reaction is robustly catalytic, such that a 1-h incubation of PC with TmpA in the presence of excess 2OG and O_2 resulted in its complete consumption (100 turnovers), as observed with TMAEP under identical conditions (Figure 6C). However, the turnover frequency was found to be less with PC (v_0/[E] = 0.074 s^{-1} at 3 °C) than with TMAEP by a factor of ~ 9. Additionally, no changes to the optical spectrum were seen upon reaction of the presumptive TmpA•Fe(II)•2OG•PC quaternary complex with O_2 (Figure S11), suggesting that the initial O_2-addition step is slow compared to subsequent events leading to re-formation of the reactant complex. In a competition experiment, TMAEP was first consumed to completion, at which point PC was then slowly consumed (Figure 6D), clearly demonstrating a preference for the phosphonate compound. Furthermore, the TmpA reaction with PC was inhibited in this competition assay compared with a reaction containing only PC (i.e., in the absence of TMAEP; compare Figure 6C, D). The inhibition was presumed to be caused by the TMAEP product, (*R*)-OH-TMAEP, which was confirmed by including it at a high concentration in a reaction with PC as the sole substrate (Figure 6C).

Structural basis for TmpA substrate specificity.—The x-ray crystal structure of the TmpA•Fe(II)•2OG complex was solved to a resolution of 1.73 Å by molecular replacement using the C-terminal catalytic domain of the *Hs* BBOX structure (PDB accession code 3O2G[75]) as the search model. As forecast by the sequence similarity that led to the misannotation of TmpA as a BBOX, the two proteins have a number of structural features in common. The overall topologies are remarkably similar (Figures 7, S12), with a root-mean-squared standard deviation (rmsd) of 2.5 Å for 355 Cα-atoms.[76] The tertiary structure of the TmpA monomer consists of an N-terminal dimerization domain (residues 1–96), a short linker (97–100), and a C-terminal catalytic domain containing the eight-stranded β-sandwich (cupin) fold characteristic of Fe/2OG oxygenases (Figure S12).[2,11,77] In both TmpA and *Hs* BBOX, the N-terminal domain facilitates assembly of a symmetric head-to-tail dimer via its interface with the C-terminal domain of the other monomer (Figure 7). Analysis by size exclusion chromatography confirmed that this quaternary structure of TmpA is also present in solution (Figure S13). The N-terminal domains of TmpA and *Hs* BBOX are structurally homologous, despite sharing < 20% sequence identity, and are composed of two three-strand antiparallel β-sheets with an intervening α/β insertion (Figures S12, S14). However, whereas *Hs* BBOX harbors an essential Zn(II) ion in the α/β insertion,[75] TmpA lacks the CxCxxC…H sequence motif required for metal binding, and a divalent cation is not observed in its N-terminal domain (Figure S14).

Structures of TmpA were also solved with the TMAEP substrate (1.70 Å resolution) and (*R*)-OH-TMAEP product (1.78 Å resolution) bound. In all three of the resulting models, the active site is located in the C-terminal domain of TmpA and contains an Fe(II) center coordinated by residues His198, Asp200, and His344 in the expected facial triad geometry (Figures 7, S15).[12,13] The dimer quaternary structure of TmpA affords two independent views of the active site in the asymmetric unit. However, in one of the two monomers (chain B), the catalytic center is occluded by crystal packing interactions (Figure S16). The lattice artifact locks an active-site lid loop into an open/disordered conformation in chain B, preventing 2OG and the primary substrate/product from binding (Figure S15). In chain A of both the TMAEP-bound and TMAEP-free structures, 2OG coordinates in bidentate mode to the iron center (Figures 7, S15) and is stabilized by a salt-bridge interaction with the side chain of conserved residue Arg352. In the absence of the primary substrate, an octahedral Fe(II) coordination geometry is completed by a water ligand axial to His344, and a sulfate ion from the crystallization solution occupies the phosphonate binding site described below (Figure S15). In the TMAEP-bound structure, the Fe(II) ion has lost the axial water ligand, presumably opening the position for dioxygen addition to initiate the reaction (Figures 7, S15). Structural comparisons show that binding of TMAEP is accompanied by inward collapse of two core beta strands (β2 and β3) that immediately precede the HxD metal-binding sequence motif (Figure S16). In addition, the loop connecting these beta strands (residues 179–196) becomes ordered and closes over the active site (described below) to shield it from solvent (Figure S16). Interestingly, this structure differs from the corresponding dynamic lid regions of many other Fe/2OG enzymes (Figure S17). This distinction is likely mandated by the location of the N-terminal domain of the other TmpA protomer, which overlaps with the typical lid-loop binding site. Closure of the unorthodox TmpA lid-loop around the TMAEP substrate results in formation of an extended water-mediated H-bonding network involving the substrate phosphonate group, the lid-loop backbone, and residues in the α/β insertion of the N-terminal domain (Figure S16). This analysis suggests that the TmpA N-terminal domain functions in substrate binding by stabilizing the closed conformation.

The active site residues that interact with TMAEP form a pocket similar to that observed in *Hs* BBOX for γbb, particularly in the vicinity of the quaternary ammonium group (Figure 7). An aromatic cage is formed by the side chains of Tyr173, Phe177, and Tyr190 around the -$N^+(CH_3)_3$ moiety (Figure 7). Tyr190 is located in the dynamic lid loop (Figure S16), consistent with a role for this structural motif in substrate binding/specificity. The faces of their aromatic rings engage in cation-π interactions, which are commonly involved in binding of substrates with quaternary ammonium groups.[78] The phosphonate dianion forms hydrogen-bonding interactions with a water molecule and the side chains of Asn187, Asn201, Tyr203, and Asn286 (Figure S16). The phosphonate group also interacts electrostatically with Arg288 (Figure 7), which is a key player in the extended H-bonding network involving the N-terminal domain of the opposite monomer (Figure S16). The fact that Arg288 is one of the few residues in the substrate-binding pocket that is not conserved in actual BBOX sequences (e.g., *Hs* Ala294) suggests that it is important in conferring specificity for the phosphonate moiety. These substrate interactions poise the C1 *pro-S* hydrogen of TMAEP above the open axial coordination site of the iron center (Figure 7), with the carbon at a distance of 4.2 Å. The product-bound structure (Figure S15) reveals clear electron density for a hydroxyl group appended to TMAEP at the C1 position with *R* stereochemistry [(*R*)-OH-TMAEP], consistent with the expected retention of configuration in the C–O-coupling step following transfer of the *pro-S* hydrogen to the ferryl complex. We presume that hydroxylation resulted here from adventitious O_2 exposure leading to turnover, during crystallization. Consistent with this interpretation, the electron density in the equatorial plane cannot be modeled well by 2OG and most likely represents a mixture of succinate (produced by turnover) and acetate (present in the synthetic TMAEP) (Figure S15). Together, the structures of TmpA in the substrate-free, substrate-bound, and product-bound states reveal dynamic changes in the protein structure that create a substrate pocket that is subtly but functionally different from that in BBOX.

TmpB is an HD-domain mixed-valent diiron oxygenase acting upon (*R*)-OH-TMAEP.

Characterization of the TmpB mixed-valent diiron cofactor.—The genomic synteny of TmpA and TmpB and the precedent provided by the PhnY/PhnZ pathway led us to anticipate that TmpB might use the product of the TmpA reaction as its substrate. This product, (*R*)-OH-TMAEP, is the trimethylammonium analogue of the PhnZ substrate, (*R*)-OH-AEP. Moreover, *Lc* TmpB shares 32% sequence identity with PhnZ (from *uncultured bacterium HF130_AEPn_1*), including conservation of its extended metal-binding sequence motif (Figure S18), suggesting that it could be a related HD-domain mixed-valent diiron oxygenase (HD-MVDO),[21,22] despite having been annotated as a phosphodiesterase.

The two known HD-MVDOs, PhnZ and MIOX, both use mixed-valent [Fe_2(II/III)] diiron cofactors to activate O_2 for four-electron oxidation reactions. We first used atomic emission spectroscopy (ICP-AES) and Mössbauer spectroscopy to determine both the nuclearity of the presumptive TmpB metallocofactor and its stable oxidation states. The His_6-tagged protein emerged from over-expression in *E. coli* and purification by immobilized metal-ion affinity chromatography with 1.0–1.3 molar equivalents of Fe and undetectable levels of other transition metals bound. TmpB enriched in ^{57}Fe (> 95 %) was incubated with chemical oxidants and reductants to probe the accessible oxidation states of the cofactor. The 120-K/0-T spectrum of the sample treated with the oxidant potassium ferricyanide (Figure 8A) is a single quadrupole doublet with parameters (isomer shift, δ, of 0.49 mm/s and quadrupole splitting parameter, ΔE_Q, of 0.84 mm/s) typical of high-spin Fe(III) ions with N/O coordination. The presence of a quadrupole doublet feature in the 4.2-K/53-mT spectrum of the same sample (Figure S19) indicates that the cofactor has an integer-spin electronic ground state, which can be rationalized by antiferromagnetic coupling of two high-spin ferric ions ($S = 5/2$) in a dinuclear cluster to give a total electron spin (S_{total}) of 0.[55] The 120-K/0-mT spectrum (Figure 8B) of the sample treated with sodium dithionite is slightly heterogenous and could be best fit as a superposition of two quadrupole doublets with parameters (δ_1 = 1.24 mm/s, ΔE_{Q1} = 3.25 mm/s, Area = 0.76; δ_2 = 1.24 mm/s, ΔE_{Q2} = 2.50 mm/s, Area = 0.23) characteristic of high-spin ferrous ions with O/N ligands, suggesting conversion of the as-isolated sample to the fully reduced Fe_2(II/II) state by the strong reductant. By contrast, treatment of the as-isolated TmpB with the milder reductant, L-ascorbate, gave a different cofactor form. Its 4.2-K/53-mT spectrum is dominated by a broad multi-line signal (Figure S20), similar to that observed for the Fe_2(II/III) complex of MIOX, which has a ground state with S_{total} = ½.[79] At higher temperature (120 K) and in the absence of an externally applied magnetic field, the paramagnetic features collapsed into quadrupole doublets (Figure S20). In this 120-K/0-T spectrum, a minor contribution (22 %) of a diferrous form was evident from the presence of a partially resolved feature matching the position of the corresponding high-energy line in the reference spectrum of the pure Fe_2(II/II) form; it was subtracted from the experimental spectrum to resolve the superimposed contributions of the other components (Figure S20). The resultant spectrum could then be accounted for as a superposition of three quadrupole doublets (Figure 8C). In the three-doublet fit, one doublet, contributing ~ 13 % of the total area, was constrained to have parameters obtained from the reference spectrum of the pure Fe_2(III/III) state. The other two variable doublets, judged to originate from the paired Fe(II) and Fe(III) ions of a mixed-valent diiron cluster, were constrained to have equal areas. This fitting analysis yielded parameters of δ_1 = 1.19 mm/s, ΔE_{Q1} = 2.57 mm/s for the high-spin Fe(II) ion and δ_2 = 0.52 mm/s, ΔE_{Q2} = 0.96 mm/s for the high-spin Fe(III) ion of the coupled mixed-valent cluster and a total contribution of 65 % of the original, experimental spectrum. The conclusion from this Mössbauer-spectroscopic analysis – that, upon treatment with an appropriately mild, one-electron reductant (ascorbate), the mixed-valent state of the diiron cluster accumulates to ~ 65 % in equilibrium with its fully reduced and fully oxidized forms – is consistent with the notion that TmpB is, like MIOX and PhnZ, an HD-MVDO.

Ascorbate-treated TmpB samples were also interrogated by EPR spectroscopy, both to verify the oxidation state of the primary cofactor form inferred by Mössbauer spectroscopy and to test for perturbations to the anticipated $g < 2$ EPR signal upon substrate binding, as seen previously for MIOX and PhnZ. In the absence of substrate (Figure 9, black), the sample exhibited a broad axial signal with apparent principal g-values < 2 [1.94, 1.78, 1.78]. This spectrum is similar to those of antiferromagnetically-coupled (S_{total} = ½) Fe_2(II/III) clusters in other proteins.[21,79,80] Comparison of the integrated intensity of this signal to that of a spin-concentration Cu(II)-EDTA standard indicated the presence of ~ 0.3 spin/iron cluster or ~ 0.6/diiron cluster, similar to the quantity deduced in the Mössbauer analysis of the parallel sample. The signal became markedly more rhombic ($\mathbf{g}$ = [1.94, 1.82 and 1.65]) upon addition of (*R*)-OH-TMAEP (Figure 9, red), suggestive of substrate binding to the cofactor, as also seen in MIOX and PhnZ.[21,79] Addition of the unmethylated analog, (*R*)-OH-AEP (the PhnZ substrate), also perturbed the EPR spectrum but gave a broader, more complex signal (Figure 9, blue). The spectrum in the presence of (*R*)-OH-AEP appears to be a superposition of more than one signal, presumably representing bound and unbound states and/or multiple, conformationally distinct forms of the (*R*)-OH-AEP complex.

TmpB-catalyzed cleavage of the C–P bond in (R)-OH-TMAEP.—With the EPR-spectroscopic data establishing that Fe_2(II/III)-TmpB binds (*R*)-OH-TMAEP under anoxic conditions, we tested for consumption of the phosphonate compound in the presence of O_2, anticipating that, by analogy to PhnZ, the C–P bond might be oxidatively cleaved to produce phosphate and the corresponding carboxylic acid [2-(trimethylammonio)acetate or glycine betaine]. Indeed, L-ascorbate-treated TmpB could consume > 200 equiv. of (*R*)-OH-TMAEP in the presence of sufficient O_2, but no conversion was observed under anoxic conditions (Figure 10A). The expected products of the reaction were verified by LC-MS (for glycine betaine) and ^{31}P NMR (for phosphate) by comparison to standards (Figure 10A). PhnZ, which natively acts on the corresponding unmethylated analog, (*R*)-OH-AEP, could also oxidize (*R*)-OH-TMAEP to yield the same products (Figure 10A). Conversely, TmpB consumed the PhnZ substrate, (*R*)-OH-AEP, very inefficiently, promoting only five turnovers (to glycine and phosphate) in a 4-h incubation (Figure 10B). This last result suggests that TmpB possesses some degree of preference for the *N*-methylated substrates, as also shown above for TmpA. To explore this preference more thoroughly, we challenged TmpB with the *N*-dimethyl analogue of its substrate, OH-DMAEP, produced from DMEAP (presumably also with *R* configuration) by TmpA in a coupled reaction. In this reaction, the expected products, phosphate and *N,N*-dimethylglycine, were readily detected by ^{31}P-NMR and LC-MS, respectively, demonstrating that TmpB can indeed oxidize OH-DMAEP (Figure S21). However, differences in the kinetics of the two hydroxylated intermediates in the context of the coupled reaction again revealed the preference of TmpB for the fully (tri)methylated substrate. Whereas (*R*)-OH-TMAEP accumulated transiently, OH-DMAEP accumulation persisted throughout the reaction as a result of its slower consumption by TmpB (Figure S21). This result demonstrates that TmpB prefers the substrate with the highest degree of *N*-methylation, as also observed for TmpA.

Structural characterization of the TmpB•Fe_2•(R)-OH-TMAEP complex.—The structure of diiron TmpB was solved by X-ray crystallography to a resolution of 1.63 Å. Phasing was achieved by molecular replacement using the PhnZ structure (PDB accession code 4MLM)[21] as the search model. The (*R*)-OH-TMAEP substrate was incorporated by soak methods; binding was observed in two of the four monomers in the asymmetric unit (chains C and D). In the other two monomers (chains A and B), lattice contacts that occlude the active site prevented substrate binding in the crystal (Figure S22). The structure therefore provides two independent views of the active site in the absence of substrate and two views of the enzyme•substrate complex. As anticipated from the 32 % sequence identity, TmpB is structurally similar to PhnZ (Figure S23),[21,24] with an rmsd of 1.5 Å for 181 Cα-atoms.[76] The conserved fold is almost entirely α-helical (Figure 11A and Figure S23). Five core helices contribute the conserved histidine and aspartate residues predicted from the extended HD-domain sequence motif (His40, His64, Asp65, His86, His109, and Asp166 of *Lc* TmpB) to coordinate a dinuclear metallocofactor, which was confirmed to be diiron by anomalous diffraction datasets collected at the Fe K-edge (Figure 11B).

In the substrate-free monomers, the active site cavity is quite solvent-exposed and the first coordination sphere is completed by a (hydr)oxo bridge and two water ligands to iron site 2 (Fe2) (Figure 11B). In place of the water ligands in the other two monomers, clear electron density for (*R*)-OH-TMAEP, continuous with the density arising from Fe2 (Figure 11C), is observed. This direct coordination of the substrate to the cofactor rationalizes the associated perturbations to its EPR spectrum. The (*R*)-OH-TMAEP molecule coordinates Fe2 in bidentate fashion via its C1 hydroxyl group and a phosphonate oxygen (Figure 11C), the same substrate binding mode previously observed in PhnZ.[21,24] The vicinal-diol substrate of MIOX (*myo*-inositol) binds in a similar chelating mode to form a five-membered ring with Fe2 of the cofactor.[25] In all three cases, the scissile C–C/P bond is opposite the Fe ion in the five-membered ring. In the TmpB structure, the phosphonate dianion is further stabilized by salt bridges with the positively-charged side chains of Arg163 and Lys113, as well as hydrogen bonds with a water molecule and the side chains of Ser131, Ser134, and Gln138 (Figure S24). The interactions contributed by this phosphonate-binding pocket are fully conserved with those in the active site of PhnZ. The C1 hydroxyl group of the TmpB substrate hydrogen bonds with the imidazole side chain of conserved His68 (Figure S24) in the sequence motif (HDIG**H**) containing the metal-coordinated histidine-aspartate dyad. The combination of (i) this hydrogen-bonding motif, (ii) the bidentate substrate coordination of Fe2, and (iii) the *R* configuration of C1 orients the lone hydrogen of C1 toward Fe1, the proposed site of dioxygen activation.[81]

In all four active site views, Fe1 has an additional ligand (Figure 11), a tyrosine residue (Tyr30) that is conserved in all TmpB-like and PhnZ-like sequences. The corresponding tyrosine (Tyr24) was also seen to coordinate Fe1 in a structure of PhnZ lacking its substrate

(Figure S23).[24] In both proteins, the Tyr ligand completes a nearly octahedral first coordination sphere of Fe1. Dioxygen binding is expected to be disfavored by this apparent coordinative saturation. In PhnZ, Tyr24 was observed not to coordinate in some structures of the PhnZ•Fe_2•(*R*)-OH-AEP complex (Figure S23), suggesting that its transient dissociation upon substrate binding opens a site for O_2 addition/activation.[24] By contrast, in our structure of substrate-soaked TmpB, the Tyr30 ligand is retained, even in the monomers with substrate bound at Fe2. This inconsistency could be a consequence of introducing the substrate by soaking of pre-formed crystals, wherein established crystal contacts could disfavor Tyr30 dissociation. Alternatively, it could have resulted from a change in the cofactor oxidation state during crystallization to either of the non-functional fully reduced or oxidized states. Regardless of the explanation for its failure to dissociate in the crystal, binding of (*R*)-OH-TMAEP by Fe_2(II/III)-TmpB in solution results in bleaching of a visible absorption feature at ~ 500 nm (Figure S25), which most likely arises from the iron-tyrosinate interaction. This observation is consistent with the hypothesis that substrate-promoted dissociation of the conserved tyrosine is functionally important in both phosphonate cleaving HD-MVDOs.

In PhnZ, substrate triggered tyrosine dissociation was proposed to be driven by interaction between the primary amine of (*R*)-OH-AEP and second-sphere Glu27, which is found in an YxxE sequence motif adjacent to the Tyr24 ligand.[24] In substrate-free PhnZ, Glu27 residue faces away from the active site (Figure S23). Conversely, in the substrate complex, both the Glu27 side chain and the secondary structural motif in which it resides undergo a large conformational change to position the Glu27 carboxylate for optimal interaction with the substrate and to induce dissociation of Tyr24 (Figure S23). In TmpB, the corresponding glutamate (Glu33) is held in place through an extensive hydrogen-bonding network in the active site (Figures S23, S24), even in the absence of substrate, and appears unlikely to be capable of the corresponding conformational change. The structure thus suggests that, if Tyr30 of TmpB does indeed dissociate upon substrate binding, the structural mechanism driving the ligand reorganization is likely to be distinct from that proposed for PhnZ.

In both TmpB and PhnZ, the YxxE loop participates in multiple substrate binding interactions. The carboxylate of the aforementioned Glu33 in TmpB is located ~3.3 Å from an *N*-methyl group of the (*R*)-OH-TMAEP quaternary amine (Figures S23, S24). At this distance, the carboxylate oxygen could be acting as an acceptor in a CH–O hydrogen-bonding interaction, made possible by the partial positive charge on the carbon of this methyl group. In addition, the backbone amide oxygen of Leu31 is within 3.5 Å of another substrate *N*-methyl group, forming a second CH–O interaction with (*R*)-OH-TMAEP (Figures S23, S24). In PhnZ, the corresponding side chain and backbone functional groups also mediate enzyme-substrate contacts, but through different intermolecular interactions. Glu27 interacts electrostatically with the (*R*)-OH-AEP amine, and the backbone amide oxygen of Ile25 (corresponding to TmpB Leu31) hydrogen bonds with what appears to be a functionally important water molecule (Figure S23). This water molecule bridges an interaction between the substrate amine, the YxxE loop, and another flexible loop that undergoes an open-to-closed transition that appears to sequester the active site from solvent upon substrate binding (Figure S23).[21,24] The analogous loop (residues 69–84) in the TmpB substrate-bound monomers remains disordered, resulting in a more exposed active site (Figures 11A, Figure S23). Failure of TmpB to undergo full active site closure upon substrate binding could again be attributed to the fact that the substrate was introduced by soaking of pre-formed crystals. Inspection of the crystal lattice contacts near the residues 69–84 loop suggests that the TmpB packing likely hinders this motion (Figure S22). The amino acid sequence of this loop contains several aromatic and acidic residues (Figure S18), which could interact favorably with the positively-charged quaternary amine of (*R*)-OH-TMAEP. However, there is no obvious aromatic cage, analogous to the one seen in TmpA, to cradle the quaternary ammonium group.

DISCUSSION

Genomically co-located genes encoding presumptive TmpA and TmpB proteins have been identified in more than 350 bacterial genomes. Automated annotation correctly classified these gene products into known structural superfamilies but failed to predict their biochemical functions.

Functional annotation of TmpA redefines substrate specificity.

The functional assignment of TmpA as a γ-butyrobetaine hydroxylase posited both the type of reaction (hydroxylation) and the primary substrate (γbb); only the former assignment was correct. Our revised assignment of TmpA as the first enzyme in a two-enzyme pathway to degrade the naturally occurring phosphonate compound, TMAEP, is based on several observations. Among all compounds screened, TMAEP was found to be the best substrate based on three metrics: binding, promotion of O_2 activation (“substrate triggering”), and overall turnover efficiency. Whereas Fe/2OG oxygenases often exhibit uncoupling of 2OG decarboxylation (to succinate) from oxidation of non-native substrates, TmpA tightly couples succinate production with TMAEP hydroxylation. TMAEP binding enhances the observed rate for reaction of the TmpA•Fe(II)•2OG cofactor with dioxygen by more than 7,500-fold. Closely related compounds either do not promote O_2 reactivity or do so to a lesser extent. O_2 activation in the presence of TMAEP leads to rapid accumulation of the canonical hydrogen-abstracting ferryl intermediate[14] with an apparent second-order rate constant of $> 1.5 \times 10^6$ $M^{-1} \cdot s^{-1}$ ($k_{1,obs} \geq 750$ s^{-1}, [**R**] ~ 0.5 mM) at 5 °C, which is, to the best of our knowledge, faster than any other Fe/2OG oxygenase for which this rate constant has been reported.[46,66,73,74] Moreover, it is known that the substrate triggering efficacies of native Fe/2OG-oxygenase substrates are generally greater than those of non-native substrates,[66] as observed for TmpA with DMAEP. Thus, the remarkably fast reaction of the TmpA•Fe(II)•2OG•TMAEP complex with O_2 is strong evidence that the *N*-trimethyl/phosphonate compound is, in fact, the native substrate. The characteristic substrate triggering in Fe/2OG enzymes is proposed to result, in part, from displacement of the axial water ligand to open up a coordination site for O_2 binding.[82,83] In the structure of the TmpA•Fe(II)•2OG•TMAEP complex, there is no density suggestive of an axial water, whereas the water ligand is clearly present when only sulfate is bound in the primary substrate site. The structural observations imply that the (potentially) general mechanism of substrate triggering – axial water dissociation – is manifested upon binding of TMAEP to TmpA.

TMAEP binding also induces structural changes beyond the first coordination sphere of the metal center that likely enable reactivity and confer substrate specificity. The contraction of the core β-barrel in the catalytic domain and closure of the lid loop positions substrate-interacting residues into the active site. This closed lid-loop conformation is stabilized by its interaction with the N-terminal domain of the opposite monomer, suggesting a role for the N-terminal domain in both dimerization and substrate binding. Although the TmpA and BBOX N-terminal domains are very similar, in terms of its secondary and tertiary structure, only BBOX harbors a Zn(II) ion in its N-terminal domain. The Zn(II) ion is coordinated by a histidine residue and three cysteine residues that are easily identified in the sequence by a **C**x**C**xx**C** motif.[75] Although the exact function of this domain and the Zn(II) ion in BBOX is unknown, removal of the zinc abolishes BBOX activity.[84] *Lc* TmpA does not possess this sequence motif, nor a Zn(II) metal in the crystal structure, yet it maintains its structure. In conjunction with genomic context information, the presence/absence of this Zn(II)-binding cysteine sequence motif could be used to differentiate TmpA from BBOX function on the basis of primary structure. The potential for this distinction to predict function of biochemically uncharacterized proteins is illustrated in the analysis of bacterial genomes that contain multiple genes annotated as BBOXs (e.g., *Halomonas* species). In the final SSN (Figure 2), the nodes representing the two proteins encoded within the same *Halomonas* genome are separated into distinct clusters, already suggestive that they perform different functions. The protein that clusters with the known *Ps* BBOX is encoded within a putative γbb degradation operon and possesses the Zn(II) binding motif (designated in the SSN by diamond nodes); it can thus be confidently assigned as a γ-butyrobetaine hydroxylase. Conversely, the protein that is clustered with *Lc* TmpA is encoded in an operon together with a TmpB-like HD domain protein and does not possess the Zn(II) binding sequence motif. We therefore suggest that it should be re-annotated as a TMAEP hydroxylase.

The occurrence of multiple genes annotated as BBOXs in a single genome indicates that BBOX and TmpA are likely paralogues, resulting from gene duplication and subsequent functional divergence. This relationship and common ancestry are evident both from their structural similarity and from the slight promiscuity of BBOX toward TMAEP. Unsurprisingly, their active sites are remarkably similar, particularly in the conserved set of aromatic residues that make cation-**π** interactions with the positively-charged quaternary ammonium common to both substrates.[85] Aromatic cages are a classic structural feature of proteins that bind substrates with quaternary ammonium groups, such as acetylcholine esterases, phosphocholine-binding antibodies, acetylcholine receptors, epigenetic trimethylysine reader domains, and trimethylamine dehydrogenases.[78] The cation-**π** interaction is primarily electrostatic, with secondary contributions from hydrophobic interactions and desolvation, which imposes a requirement for a positively-charged functionality for efficient substrate recognition and binding.[86,87] This structurally-induced discrimination is obvious in the lack of TmpA activity toward phosphonates lacking any amine functionality. More subtly, TmpA even discriminates further on the basis of the degree of *N*-methylation, despite the fact that, at pH 7, the primary and tertiary amine groups of both 2-AEP and DMAEP should be protonated/cationic. Similar discrimination is observed for epigenetic methyllysine reader domains[88,89] and alkylamine dehydrogenases.[78] The reader domains that bind substrates with higher degrees of lysine *N*-methylation possess pockets with up to four aromatic residues, similar to that observed in TmpA. Conversely, reader domains that recognize lower degrees of lysine *N*-methylation substitute one or more of the aromatic residues with residues capable of hydrogen bonding with the extra proton(s) on the amine. Furthermore, the *N*-methyl groups harbor some of the distributed positive charge from the nitrogen,[86] perhaps enhancing the binding interaction and conferring the observed specificity of TmpA for higher degrees of *N*-methylation.

The quaternary ammonium group is not, however, the sole basis for substrate discrimination by TmpA, as the phosphonate functionality is also essential. It is also the major distinguishing feature between the substrates of TmpA and BBOX. A simple but crucial difference between the TmpA and BBOX active sites is the presence of Arg288 in TmpA, which interacts electrostatically with the phosphonate moiety of TMAEP. Conversely, an alanine residue is found at this position in *Hs* BBOX. A survey of reported substrate-bound structures of enzymes that utilize phosphonate substrates reveals that there is almost invariably at least one Arg that interacts with the phosphonate functional group (e.g., phosphonoacetaldehyde dehydrogenase,[90] DhpI,[91] 2-AEP transaminase,[92] PhnZ,[21,24] and HppE[93]). This list now also includes TmpB, which has a conserved arginine residue (Arg163) interacting with the phosphonate moiety of (*R*)-OH-TMAEP. The specificity of TmpA for the phosphonate functional group might suggest *a priori* that it would share structural homology and evolutionary history with the organophosphonate hydroxylase, PhnY. Indeed, they enact identical chemical transformations – hydroxylation of a carbon adjacent to a phosphonate – on substrates that differ only in amine methylation. Yet TmpA and PhnY are only modestly similar in sequence, and PhnY lacks the N-terminal domain common to both TmpA and BBOX. Instead, it is the regioselectivity of hydroxylation with respect to the quaternary ammonium that is common to TmpA and its homologue, BBOX. Thus, TmpA and PhnY likely evolved convergently to promote similar reactions that initiate phosphonate degradation, to be completed by their HD-MVDO partners.

Functional annotation of TmpB redefines reaction and substrate specificity.

Conversely, these partners, TmpB and PhnZ, likely share an evolutionary history that is reflected by their cofactor, the chemical transformation they promote (oxidative C–P cleavage), and their structural similarity. The x-ray crystallographic and spectroscopic (Mössbauer and EPR) data establish that TmpB harbors a diiron cofactor. Importantly, the data show that TmpB shares with HD-MVDOs the ability to stabilize the mixed-valent Fe_2(II/III) oxidation state to as much as 60–70 % in equilibrium with the more oxidized and reduced cofactor forms and to deploy this state as a cofactor.[21,23] The longer-known ferritin-like diiron oxygenases react with dioxygen from the Fe_2(II/II) oxidation state and can typically accumulate the mixed-valent form to only modest fractions (< 10 %).[94–102] Although the prediction of HD-MVDOs cannot rely solely on identification of the extended sequence motif of metal ligands, the proteins that possess this motif could serve as a pool to be mined for novel oxygenase functions. Thus, the metallocofactor nuclearity should be considered in functional annotation of existing and newly identified HD-domain proteins.

The substrate preference of TmpB for its *N*-trimethylated substrate mirrors the specificity of TmpA. However, the structural basis for this discrimination by TmpB is not immediately obvious from a simple comparison of its structure to that of PhnZ. The primary amine of (*R*)-OH-AEP forms a single contact with the PhnZ protein via the side chain of Glu27, which is also conserved in TmpB (Glu33). The fact that the other interactions with the primary amine in PhnZ are mediated by water molecules may explain its greater substrate promiscuity. Conversely, the preference of TmpB for substrates with higher degrees of *N*-methylation is achieved by a more well-defined binding pocket mediated by direct protein contacts. In the structure of the TmpB•(*R*)-OH-TMAEP complex, the quaternary ammonium ion establishes CH–O hydrogen bonds[103,104] with the protein via its *N*-methyl groups, which harbor some positive charge.[86] Although there is more precedent for cation-π interactions of quaternary-ammonium-containing substrates,[78] CH–O interactions have been observed in the binding pocket of the choline TMA-lyase enzyme CutC[105] and have been invoked in protein interactions with the sulfonium methyl group of *S*-adenosylmethionine.[106] In addition to these observed substrate contacts, the lid loop containing residues 69–84 in TmpB, which diverges greatly in sequence from that of PhnZ, likely plays a key role in substrate recognition. Unfortunately, due to the soaking method used to introduce the substrate and the obstructive crystal contacts, the lid loop remained disordered, preventing validation of this hypothesis.

The observed substrate promiscuity of PhnZ for *N*-methylated substrates is not expected to be relevant in nature, because its partner hydroxylase, PhnY, is unable to convert the *N*-methylated substrate into the hydroxylated intermediate for PhnZ to utilize. Thus, for both the TmpA/B and the PhnY/Z pair, the specificity for degradation of the closely related phosphonate compounds is determined primarily by the Fe/2OG dioxygenase, rather than the HD-MVDO partner. Interestingly, we identified 16 cases in the sequence databases of a single operon possessing a single HD-domain protein but both BBOX-like and PhnY-like (PhyH-like) Fe/2OG oxygenases. While this type of operon could potentially represent a novel degradation pathway, it might also reflect the specialization of the Fe/2OG oxygenases and contrasting versatility of the HD-MVDOs. The Fe/2OG oxygenases could process different substrates, but the HD-MVDO could degrade the products of both. The apparent dedication of the pathway partners – an Fe/2OG dioxygenase and an HD-MVDO – is also reflected by the co-occurrence of their encoding genes in a single operon. In addition to the more than 300 examples of operons containing separate genes for TmpA and TmpB, there are at least 59 examples of genes that appear to encode fusions of TmpA and TmpB domains. There are also extant gene fusions of PhnY and PhnZ. It is unknown whether these fusions fold to be functional two-domain enzymes, but there is precedent for an active bifunctional protein with fused Fe/2OG oxygenase and methyltransferase domains.[107] Nevertheless, the occurrence of the gene fusions supports the notions that each pairing is specific and that the pair functions together.[108] Furthermore, the frequent presence of a gene encoding a LysR-type transcriptional regulator upstream of the TmpA/B genes suggests that they are co-transcribed and is consistent with the biochemical evidence presented herein that their gene products constitute a degradation pathway. Similarly, transcription of the operon encoding the PhnY/Z pair appears to be under control of a LysR-type regulator. Transcriptional regulators of the LysR family[109] are commonly found upstream of other known phosphonate degradation operons and, in some cases, have been shown to be induced by the specific phosphonate compounds that the gene operon products catabolize.[110–112]

These precedents suggest that transcription of the appropriate degradation pair, TmpA/B or PhnY/Z, would be specifically induced by the presence of TMAEP or 2-AEP, respectively.

Relevance of TmpA/B degradation pathway in microbes.

Phosphonates and their degradation products serve valuable roles for the bacteria that occupy certain environmental niches. Phosphonates exist in nature as components of phosphonolipids, phosphonoproteins, and phosphonoglycans and are incorporated in the place of their phosphoester analogues, presumably to confer the advantage of a more stable C–P bond in place of the hydrolyzable C–O–P bond.[113–115] Of the most common phosphoesters, phosphoethanolamine, phosphoserine, and phosphocholine, each has a phosphonate counterpart (2-AEP, phosphonoalanine, and TMAEP, respectively). While 2-AEP is the most abundant phosphonate in nature, TMAEP, as well as DMAEP and *N*-methyl-AEP, have been reported in protozoa and marine invertebrates such as sea anemones, marine plankton, and algae.[114,116–118] Degradation pathways for both 2-AEP and phosphonoalanine have been discovered, ultimately producing phosphate and a source of carbon.[119–123] The TmpA/B pathway mirrors the PhnY/Z degradation pathway for 2-AEP in its chemical transformations but is instead specific for TMAEP. It represents the first example of a pathway for microbial metabolism of this naturally occurring phosphonate compound.

The TmpA/B operon is found primarily in Proteobacteria and Actinobacteria, with considerable representation in marine and N_2-fixing plant-symbiotic bacteria. The products of this degradation pathway, phosphate and glycine betaine, are valuable nutrients for such bacteria. Organophosphonates have been established as major sources of phosphorus in marine environments, which are limited in this essential nutrient.[119] Phosphorus has been found to enhance plant growth, promote formation of crucial root nodules where the symbiotic bacteria reside, and support higher levels of N_2-fixation by the symbiotes.[124–128] Many bacteria possess multiple phosphonate degradation pathways to extract phosphate, but the requirement for different specialized pathways and their cooperation are poorly understood.[119–122] Nevertheless, the ubiquity of mechanisms for the release of phosphate from dissolved organic phosphorus sources highlights their necessity and utility. The second product of the TmpA/B pathway, glycine betaine, is a known osmoprotectant. In marine bacteria, osmoregulation by compounds with quaternary ammonium functionality is common to protect against high salinity.[129] Glycine betaine is the most potent osmoprotectant in some N_2-fixing bacteria, enhancing their growth under conditions of osmotic stress,[131,132] which diminishes their ability to fix nitrogen.[130,131,132] Glycine betaine can also be utilized as a source of carbon and nitrogen upon its demethylation to glycine.[133–135] Degradation of TMEAP for nutrients could provide a competitive advantage for soil bacteria,[134] particularly for colonization of the environmental niche of plant roots.

CONCLUSIONS

We have defined the biochemical reactions of the Fe/2OG oxygenase, TmpA, and the HD-domain protein, TmpB, to elucidate a new pathway for degradation of the natural organophosphonate, TMAEP. The combination of genomic context and structure-informed sequence motifs will allow for the proper identification and annotation of the genes associated with this pathway in existing and newly sequenced genomes. More broadly, the biochemical and structural information gleaned about these subsets of proteins could potentially inform bioinformatic analyses aiming to categorize and predict functions within their large superfamilies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank Prof. Christopher J. Schofield (University of Oxford, UK) for the library of betaine compounds and *Ps* BBOX plasmid, Prof. Wilfred A. van der Donk (University of Illinois, Urbana-Champaign) for the library of phosphonate compounds, the Shared Fermentation Facility of The Pennsylvania State University (University Park, PA) for use of the Microfluidics M-110EH-30 microfluidizer processor, Mr. Henry Gong at the Pennsylvania State Materials Research Institute (University Park, PA) for ICP-AES analysis, and Dr. Emmanuel Hatzakis at the Pennsylvania State University Nuclear Magnetic Resonance facility.

Funding. This work was funded by the National Science Foundation (MCB-1330784 to J.M.B. and C.K.) and the National Institutes of Health (GM119707 to A.K.B, GM111978 to M.-E.P., and GM127079 to C.K.). Use of the Advanced Photon Source, an Office of Science User Facility operated for the US Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the US DOE under Contract DE-AC02-06CH11357. GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and Michigan Technology Tri-Corridor Grant 085P1000817.

Abbreviations:

Fe/2OG	iron(II)- and 2-oxoglutarate-dependent
MIOX	*myo*-inositol oxygenase
2-AEP	2-aminoethylphosphonate
(*R*)-OH-AEP	(R)-1-hydroxy-2-aminoethylphosphonate
BBOX	γ-butyrobetaine hydroxylase
γbb	γ-butyrobetaine
Lc	*Leisingera caerulea*
TMAEP	2-(trimethylammonio)ethylphosphonate
(*R*)-OH-TMAEP	(R)-1-hydroxy-2-(trimethylammonio)ethylphosphonate
Ps	*Pseudomonas sp. AK-1*
Hs	*Homo sapiens*
PC	phosphocholine
GABA	γ-aminobutyric acid

γbb-3	3-(trimethylammonio)propionic acid
TMLH	ε-trimethyllysine hydroxylase
DMAEP	2-(dimethylamino)ethylphosphonate
HD-MVDO	HD-domain mixed-valent diiron oxygenase

REFERENCES

[1]. Sono M, Roach MP, Coulter ED, Dawson JH (1996) Heme-containing oxygenases. Chem. Rev 96, 2841–2888. [PubMed: 11848843]

[2]. Hausinger RP (2015) 2-Oxoglutarate-Dependent Oxygenases, Hausinger RP, Schofield CJ, Ed., The Royal Society of Chemistry.

[3]. Andrews SC (2010) The ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. Biochim. Biophys. Acta 1800, 691–705. [PubMed: 20553812]

[4]. Aravind L, Koonin EV (1998) The HD domain defines a new superfamily of metal-dependent phosphohydrolases. Trends Biochem. Sci 23, 469–472. [PubMed: 9868367]

[5]. Baier F, Copp JN, Tokuriki N (2016) Evolution of enzyme superfamilies: Comprehensive exploration of sequence-function relationships. Biochemistry 55, 6375–6388. [PubMed: 27802036]

[6]. Roche DB, Bruls T (2015) The enzymatic nature of an anonymous protein sequence cannot reliably be inferred from superfamily level structural information alone. Protein Sci 24, 643–650. [PubMed: 25559918]

[7]. Galperin MY, Koonin EV (2012) Divergence and convergence in enzyme evolution. J. Biol. Chem 287, 21–28. [PubMed: 22069324]

[8]. Gaudet P, Livstone MS, Lewis SE, Thomas PD (2011) Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. Brief. Bioinform 12, 449–462. [PubMed: 21873635]

[9]. du Plessis L, Skunca N, Dessimoz C (2011) The what, where, how and why of gene ontology--a primer for bioinformaticians. Brief Bioinform 12, 723–735. [PubMed: 21330331]

[10]. Skunca N, Altenhoff A, Dessimoz C (2012) Quality of computationally inferred gene ontology annotations. PLoS Comput. Biol 8, e1002533. [PubMed: 22693439]

[11]. Aik W, McDonough MA, Thalhammer A, Chowdhury R, Schofield CJ (2012) Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. Curr. Opin. Struct. Biol 22, 691–700. [PubMed: 23142576]

[12]. Koehntop KD, Emerson JP, Que L Jr. (2005) The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes. J. Biol. Inorg. Chem 10, 87–93. [PubMed: 15739104]

[13]. Hegg EL, Que L Jr. (1997) The 2-His-1-carboxylate facial triad - An emerging structural motif in mononuclear non-heme iron(II) enzymes. Eur. J. Biochem 250, 625–629. [PubMed: 9461283]

[14]. Price JC, Barr EW, Tirupati B, Bollinger JM Jr., Krebs C (2003) The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: A high-spin Fe(IV) complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. Biochemistry 42, 7497–7508. [PubMed: 12809506]

[15]. Bollinger JM Jr., Chang W.-c., Matthews ML, Martinie RJ, Boal AK, Krebs C (2015) Mechanisms of 2-oxoglutarate-dependent oxygenases: The hydroxylation paradigm and beyond, In 2-Oxoglutarate-dependent oxygenases, pp 95–122, The Royal Society of Chemistry.

[16]. Murray KD, Bremer H (1996) Control of spoT-dependent ppGpp synthesis and degradation in *Escherichia coli*. J. Mol. Biol 259, 41–57. [PubMed: 8648647]

[17]. Francis SH, Colbran JL, McAllister-Lucas LM, Corbin JD (1994) Zinc interactions and conserved motifs of the cGMP-binding cGMP-specific phosphodiesterase suggest that it is a zinc hydrolase. J. Biol. Chem 269, 22477–22480. [PubMed: 8077192]

[18]. Johnson GS, Adler CR, Collins JJ, Court D (1979) Role of the spoT gene product and manganese ion in the metabolism of guanosine 5'-diphosphate 3'-diphosphate in *Escherichia coli*. J. Biol. Chem 254, 5483–5487. [PubMed: 376509]

[19]. Quirk S, Bessman MJ (1991) dGTP triphosphohydrolase, a unique enzyme confined to members of the family *Enterobacteriaceae*. J. Bacteriol 173, 6665–6669. [PubMed: 1657865]

[20]. Bollinger JM Jr., Diao Y, Matthews ML, Xing G, Krebs C (2009) *myo*-Inositol oxygenase: a radical new pathway for O(2) and C-H activation at a nonheme diiron cluster. Dalton Trans, 905–914. [PubMed: 19173070]

[21]. Wörsdörfer B, Lingaraju M, Yennawar NH, Boal AK, Krebs C, Bollinger JM Jr., Pandelia M-E (2013) Organophosphonate-degrading PhnZ reveals an emerging family of HD domain mixed-valent diiron oxygenases. Proc. Natl. Acad. Sci. U S A 110, 18874–18879. [PubMed: 24198335]

[22]. McSorley FR, Wyatt PB, Martinez A, DeLong EF, Hove-Jensen B, Zechel DL (2012) PhnY and PhnZ comprise a new oxidative pathway for enzymatic cleavage of a carbon-phosphorus bond. J. Am. Chem. Soc 134, 8364–8367. [PubMed: 22564006]

[23]. Xing G, Barr EW, Diao Y, Hoffart LM, Prabhu KS, Arner RJ, Reddy CC, Krebs C, Bollinger JM Jr. (2006) Oxygen activation by a mixed-valent, diiron(II/III) cluster in the glycol cleavage reaction catalyzed by myo-inositol oxygenase. Biochemistry 45, 5402–5412. [PubMed: 16634621]

[24]. van Staalduinen LM, McSorley FR, Schiessl K, Seguin J, Wyatt PB, Hammerschmidt F, Zechel DL, Jia Z (2014) Crystal structure of PhnZ in complex with substrate reveals a di-iron oxygenase mechanism for catabolism of organophosphonates. Proc. Natl. Acad. Sci. U S A 111, 5171–5176. [PubMed: 24706911]

[25]. Brown PM, Caradoc-Davies TT, Dickson JM, Cooper GJ, Loomes KM, Baker EN (2006) Crystal structure of a substrate complex of myo-inositol oxygenase, a di-iron oxygenase with a key role in inositol metabolism. Proc. Natl. Acad. Sci. U S A 103, 15032–15037. [PubMed: 17012379]

[26]. Lovering AL, Capeness MJ, Lambert C, Hobley L, Sockett RE (2011) The structure of an unconventional HD-GYP protein from Bdellovibrio reveals the roles of conserved residues in this class of cyclic-di-GMP phosphodiesterases. mBio 2, e00163–00111. [PubMed: 21990613]

[27]. Benda C, Ebert J, Scheltema RA, Schiller HB, Baumgartner M, Bonneau F, Mann M, Conti E (2014) Structural model of a CRISPR RNA-silencing complex reveals the RNA-target cleavage activity in Cmr4. Mol. Cell 56, 43–54. [PubMed: 25280103]

[28]. Jung TY, An Y, Park KH, Lee MH, Oh BH, Woo E (2015) Crystal structure of the Csm1 subunit of the Csm complex and its single-stranded DNA-specific nuclease activity. Structure 23, 782–790. [PubMed: 25773141]

[29]. Gerlt JA, Allen KN, Almo SC, Armstrong RN, Babbitt PC, Cronan JE, Dunaway-Mariano D, Imker HJ, Jacobson MP, Minor W, Poulter CD, Raushel FM, Sali A, Shoichet BK, Sweedler JV (2011) The Enzyme Function Initiative. Biochemistry 50, 9950–9962. [PubMed: 21999478]

[30]. Peng J, Uygun S, Kim T, Wang Y, Rhee SY, Chen J (2015) Measuring semantic similarities by combining gene ontology annotations and gene co-function networks. BMC Bioinform 16.

[31]. Sjölander K, Datta RS, Shen Y, Shoffner GM (2011) Ortholog identification in the presence of domain architecture rearrangement. Brief. Bioinform 12, 413–422. [PubMed: 21712343]

[32]. Teng Z, Guo M, Liu X, Dai Q, Wang C, Xuan P (2013) Measuring gene functional similarity based on group-wise comparison of GO terms. Bioinformatics 29, 1424–1432. [PubMed: 23572412]

[33]. Wolf YI, Rogozin IB, Kondrashov AS, Koonin EV (2001) Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. Genome Res 11, 356–372. [PubMed: 11230160]

[34]. Korbel JO, Jensen LJ, von Mering C, Bork P (2004) Analysis of genomic context: prediction of functional associations from conserved bidirectionally transcribed gene pairs. Nat. Biotechnol 22, 911–917. [PubMed: 15229555]

[35]. Rogozin IB, Makarova KS, Murvai J, Czabarka E, Wolf YI, Tatusov RL, Szekely LA, Koonin EV (2002) Connected gene neighborhoods in prokaryotic genomes. Nucleic Acids Res 30, 2212–2223. [PubMed: 12000841]

[36]. Gabaldon T, Huynen MA (2004) Prediction of protein function and pathways in the genome era. Cell. Mol. Life Sci 61, 930–944. [PubMed: 15095013]

[37]. Huynen M, Snel B, Lath WI, Bork P (2000) Exploitation of gene context. Curr. Op. Struct. Biol 10, 366–370.

[38]. Bremer J (1983) Carnitine--metabolism and functions. Physiol. Rev 63, 1420–1480. [PubMed: 6361812]

[39]. Lindstedt G, Lindstedt S, Midtvedt T, Tofft M (1970) Inducible gamma-butyrobetaine-degrading enzymes in *Pseudomonas* species AK 1. J. Bacteriol 101, 1094–1095. [PubMed: 5438040]

[40]. Rebouche CJ, Seim H (1998) Carnitine metabolism and its regulation in microorganisms and mammals. Annu. Rev. Nutr 18, 39–61. [PubMed: 9706218]

[41]. Vandecandelaere I, Segaert E, Mollica A, Faimali M, Vandamme P (2008) *Leisingera aquimarina* sp. nov., isolated from a marine electroactive biofilm, and emended descriptions of *Leisingera methylohalidivorans* Schaefer et al. 2002, *Phaeobacter daeponensis* Yoon et al. 2007 and *Phaeobacter inhibens* Martens et al. 2006. Int. J. Syst. Evol. Microbiol 58, 2788–2793. [PubMed: 19060059]

[42]. Breider S, Scheuner C, Schumann P, Fiebig A, Petersen J, Pradella S, Klenk HP, Brinkhoff T, Goker M (2014) Genome-scale data suggest reclassifications in the *Leisingera-Phaeobacter* cluster including proposals for *Sedimentitalea* gen. nov. and *Pseudophaeobacter* gen. nov. Front. Microbiol 5, 416. [PubMed: 25157246]

[43]. Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, Whalen KL (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. BBA-Proteins Proteom 1854, 1019–1037.

[44]. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498–2504. [PubMed: 14597658]

[45]. Markowitz VM, Chen IMA, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang JH, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC (2012) IMG: the integrated microbial genomes database and comparative analysis system. Nucleic Acids Res 40, D115–D122. [PubMed: 22194640]

[46]. Price JC, Barr EW, Hoffart LM, Krebs C, Bollinger JM J. (2005) Kinetic dissection of the catalytic mechanism of taurine:α-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. Biochemistry 44, 8138–8147. [PubMed: 15924433]

[47]. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. Method. Enzymol 276, 307–326.

[48]. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta. Crystallogr. D Biol. Crystallogr 67, 355–367. [PubMed: 21460454]

[49]. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta. Crystallogr. D Biol. Crystallogr 66, 486–501. [PubMed: 20383002]

[50]. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta. Crystallogr. D Biol. Crystallogr 66, 12–21. [PubMed: 20057044]

[51]. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. J. Appl. Crystallogr 40, 658–674. [PubMed: 19461840]

[52]. Langer GG, Hazledine S, Wiegels T, Carolan C, Lamzin VS (2013) Visual automated macromolecular model building. Acta. Crystallogr. D Biol. Crystallogr 69, 635–641. [PubMed: 23519672]

[53]. Lebedev AA, Young P, Isupov MN, Moroz OV, Vagin AA, Murshudov GN (2012) JLigand: a graphical tool for the CCP4 template-restraint library. Acta. Crystallogr. D Biol. Crystallogr 68, 431–440. [PubMed: 22505263]

[54]. Vagin A, Teplyakov A (1997) MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr 30, 1022–1025.

[55]. The 4.2 K/53 mT spectrum also shows broad baseline features in addition to the quadrupole doublet, suggesting the presence of a mononuclear high spin ferric (S = 5/2) species. This minority species, which is also detected by EPR spectroscopy with an effective g-value of ~ 4.3, is attributed to adventitiously-bound iron observed in the x-ray crystal structure, rather than partial occupancy of the diiron cofactor (Figure S19)

[56]. Immirzi A (1966), Ahmed FR, Ed., Munksgaard.

[57]. Ten Eyck LF (1973) Crystallographic Fast Fourier Transforms. Acta Cryst A29, 183–191.

[58]. Read RJ, Schierbeek AJ (1988) A phased translation function. J. Appl. Cryst 21, 490–495.

[59]. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr 67, 235–242. [PubMed: 21460441]

[60]. Lindstedt G, Linstedt S, Midtvedt T, Tofft M (1967) Hydroxylation of gamma-butyrobetaine to carnitine. Biochem. J 103, 1271–1282.

[61]. Lindstedt G, Lindstedt S, Olander B, Tofft M (1968) Alpha-ketoglutarate and hydroxylation of gamma-butyrobetaine. Biochim. Biophys. Acta 158, 503–505. [PubMed: 5660119]

[62]. Lindstedt G, Lindstedt S, Nordin I (1977) Purification and Properties of Gamma-Butyrobetaine Hydroxylase from *Pseudomonas-Sp*-Ak-1. Biochemistry 16, 2181–2188. [PubMed: 861203]

[63]. Atkinson HJ, Morris JH, Ferrin TE, Babbitt PC (2009) Using Sequence Similarity Networks for visualization of relationships across diverse protein superfamilies. PLoS One 4, e4345. [PubMed: 19190775]

[64]. Kazaks A, Makrecka-Kuka M, Kuka J, Voronkova T, Akopjana I, Grinberga S, Pugovics O, Tars K (2014) Expression and purification of active, stabilized trimethyllysine hydroxylase. Protein Expr. Purif 104, 1–6. [PubMed: 25220864]

[65]. Pavel EG, Zhou J, Busby RW, Gunsior M, Townsend CA, Solomon EI (1998) Circular dichroism and magnetic circular dichroism spectroscopic studies of the non-heme ferrous active site in clavaminate synthase and its interaction with alpha-ketoglutarate cosubstrate. J. Am. Chem. Soc 120, 743–753.

[66]. Matthews ML, Krest CM, Barr EW, Vaillancourt FH, Walsh CT, Green MT, Krebs C, Bollinger JM Jr. (2009) Substrate-triggered formation and remarkable stability of the C-H bond-cleaving chloroferryl intermediate in the aliphatic halogenase, SyrB2. Biochemistry 48, 4331–4343. [PubMed: 19245217]

[67]. Ryle MJ, Padmakumar R, Hausinger RP (1999) Stopped-flow kinetic analysis of Escherichia coli taurine/alpha-ketoglutarate dioxygenase: Interactions with alpha-ketoglutarate, taurine, and oxygen. Biochemistry 38, 15278–15286. [PubMed: 10563813]

[68]. Solomon EI, Brunold TC, Davis MI, Kemsley JN, Lee SK, Lehnert N, Neese F, Skulan AJ, Yang YS, Zhou J (2000) Geometric and electronic structure/function correlations in non-heme iron enzymes. Chem. Rev 100, 235–349. [PubMed: 11749238]

[69]. Decker A, Clay MD, Solomon EI (2006) Spectroscopy and electronic structures of mono- and binuclear high-valent non-heme iron-oxo systems. J. Inorg. Biochem 100, 697–706. [PubMed: 16510189]

[70]. Krebs C, Fujimori DG, Walsh CT, Bollinger JM Jr. (2007) Non-heme Fe(IV)-oxo intermediates. Acc. Chem. Res 40, 484–492. [PubMed: 17542550]

[71]. Hoffart LM, Barr EW, Guyer RB, Bollinger JM Jr, Krebs C (2006) Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. Proc. Natl. Acad. Sci. U S A 103, 14738–14743. [PubMed: 17003127]

[72]. Pan J, Bhardwaj M, Zhang B, Chang W. c., Schardl CL, Krebs C, Grossman RB, Bollinger JM Jr. (2018) Installation of the ether bridge of lolines by the iron- and 2-oxoglutarate-dependent oxygenase, LolO: Regio- and stereochemistry of sequential hydroxylation and oxacyclization reactions. Biochemistry 57, 2074–2083. [PubMed: 29537853]

[73]. Dunham NP, Chang W. c., Mitchell AJ, Martinie RJ, Zhang B, Bergman JA, Rajakovich LJ, Wang B, Silakov A, Krebs C, Boal AK, Bollinger JM Jr. (2018) Two distinct mechanisms for C-C desaturation by iron(II)- and 2-(oxo)glutarate-dependent oxygenases: Importance of alpha-heteroatom assistance. J. Am. Chem. Soc 140, 7116–7126. [PubMed: 29708749]

[74]. Dunham NP, Mitchell AJ, Del Rio Pantoja JM, Krebs C, Bollinger JM Jr., Boal AK (2018) α-Amine desaturation of d-arginine by the iron(II)- and 2-(oxo)glutarate-dependent l-arginine 3-hydroxylase, VioC. Biochemistry 57, 6479–6488. [PubMed: 30403469]

[75]. Leung IK, Krojer TJ, Kochan GT, Henry L, von Delft F, Claridge TD, Oppermann U, McDonough MA, Schofield CJ (2010) Structural and mechanistic studies on gamma-butyrobetaine hydroxylase. Chem. Biol 17, 1316–1324. [PubMed: 21168767]

[76]. Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta. Crystallogr. D Biol. Crystallogr 60, 2256–2268. [PubMed: 15572779]

[77]. Dunwell JM, Purvis A, Khuri S (2004) Cupins: the most functionally diverse protein superfamily? Phytochemistry 65, 7–17. [PubMed: 14697267]

[78]. Ma JC, Dougherty DA (1997) The cation-pi interaction. Chem. Rev 97, 1303–1324. [PubMed: 11851453]

[79]. Xing G, Hoffart LM, Diao Y, Prabhu KS, Arner RJ, Reddy CC, Krebs C, Bollinger JM Jr. (2006) A coupled dinuclear iron cluster that is perturbed by substrate binding in myo-inositol oxygenase. Biochemistry 45, 5393–5401. [PubMed: 16634620]

[80]. Doi K, McCracken J, Peisach J, Aisen P (1988) The binding of molybdate to uteroferrin. Hyperfine interactions of the binuclear center with 95Mo, 1H, and 2H. J. Biol. Chem 263, 5757–5763. [PubMed: 2833515]

[81]. Hirao H, Morokuma K (2009) Insights into the (superoxo)Fe(III)Fe(III) intermediate and reaction mechanism of myo-inositol oxygenase: DFT and ONIOM(DFT:MM) study. J. Am. Chem. Soc 131, 17206–17214. [PubMed: 19929019]

[82]. Zhou J, Kelly WL, Bachmann BO, Gunsior M, Townsend CA, Solomon EI (2001) Spectroscopic studies of substrate interactions with clavaminate synthase 2, a multifunctional alpha-KG-dependent non-heme iron enzyme: Correlation with mechanisms and reactivities. J. Am. Chem. Soc 123, 7388–7398. [PubMed: 11472170]

[83]. Neidig ML, Brown CD, Light KM, Fujimori DG, Nolan EM, Price JC, Barr EW, Bollinger JM Jr., Krebs C, Walsh CT, Solomon EI (2007) CD and MCD of CytC3 and taurine dioxygenase: Role of the facial triad in alpha-KG-dependent oxygenases. J. Am. Chem. Soc 129, 14224–14231. [PubMed: 17967013]

[84]. Rydzik AM, Brem J, Struwe WB, Kochan GT, Benesch JL, Schofield CJ (2014) Ejection of structural zinc leads to inhibition of gamma-butyrobetaine hydroxylase. Bioorg. Med. Chem. Lett 24, 4954–4957. [PubMed: 25266780]

[85]. Kamps JJ, Khan A, Choi H, Lesniak RK, Brem J, Rydzik AM, McDonough MA, Schofield CJ, Claridge TD, Mecinovic J (2016) Cation-pi interactions contribute to substrate recognition in gamma-butyrobetaine hydroxylase catalysis. Chemistry 22, 1270–1276. [PubMed: 26660433]

[86]. Mecozzi S, West AP, Dougherty DA (1996) Cation-pi interactions in aromatics of biological and medicinal interest: Electrostatic potential surfaces as a useful qualitative guide. Proc. Natl. Acad. Sci. USA 93, 10566–10571. [PubMed: 8855218]

[87]. Mecozzi S, West AP, Dougherty DA (1996) Cation-pi interactions in simple aromatics: Electrostatics provide a predictive tool. J. Am. Chem. Soc 118, 2307–2308.

[88]. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat. Struct. Mol. Biol 14, 1025–1040. [PubMed: 17984965]

[89]. Kamps JJ, Huang J, Poater J, Xu C, Pieters BJ, Dong A, Min J, Sherman W, Beuming T, Matthias Bickelhaupt F, Li H, Mecinovic J (2015) Chemical basis for the recognition of trimethyllysine by epigenetic reader proteins. Nat. Commun 6, 8911. [PubMed: 26578293]

[90]. Agarwal V, Peck SC, Chen JH, Borisova SA, Chekan JR, van der Donk WA, Nair SK (2014) Structure and function of phosphonoacetaldehyde dehydrogenase: The missing link in phosphonoacetate formation. Chem. Biol 21, 125–135. [PubMed: 24361046]

[91]. Lee JH, Bae B, Kuemin M, Circello BT, Metcalf WW, Nair SK, van der Donk WA (2010) Characterization and structure of DhpI, a phosphonate O-methyltransferase involved in dehydrophos biosynthesis. Proc. Natl. Acad. Sci. U. S. A 107, 17557–17562. [PubMed: 20876132]

[92]. Chen CCH, Zhang H, Kim AD, Howard A, Sheldrick GM, Mariano-Dunaway D, Herzberg O (2002) Degradation pathway of the phosphonate ciliatine: Crystal structure of 2-aminoethylphosphonate transaminase. Biochemistry 41, 13162–13169. [PubMed: 12403617]

[93]. Yun D, Dey M, Higgins LJ, Yan F, Liu HW, Drennan CL (2011) Structural basis of regiospecificity of a mononuclear iron enzyme in antibiotic fosfomycin biosynthesis. J. Am. Chem. Soc 133, 11262–11269. [PubMed: 21682308]

[94]. Fox BG, Hendrich MP, Surerus KK, Andersson KK, Froland WA, Lipscomb JD, Münck E (1993) Mössbauer, EPR, and ENDOR studies of the hydroxylase and reductase components of methane monooxygenase from *Methylosinus trichosporium* Ob3b. J. Am. Chem. Soc 115, 3688–3701.

[95]. Woodland MP, Patil DS, Cammack R, Dalton H (1986) Electron-spin-resonance studies of protein-A of the soluble methane monooxygenase from Methylococcus capsulatus (Bath). Biochim. Biophys. Acta 873, 237–242.

[96]. Fox BG, Surerus KK, Münck E, Lipscomb JD (1988) Evidence for a μ-oxo-bridged binuclear iron cluster in the hydroxylase component of methane monooxygenase - Mössbauer and electron-paramagnetic-resonance studies. J Biol. Chem 263, 10553–10556. [PubMed: 2839495]

[97]. Han JY, Swarts JC, Sykes AG (1996) Kinetic studies on the hydrazine and phenylhydrazine reductions of the *Escherichia coli* R2 subunit of ribonucleotide reductase. Inorg. Chem 35, 4629–4634.

[98]. Davydov R, Kuprin S, Gräslund A, Ehrenberg A (1994) Electron-paramagnetic-resonance study of the mixed-valent diiron center in Escherichia coli ribonucleotide reductase produced by reduction of radical-free protein-R2 at 77-K. J. Am. Chem. Soc 116, 11120–11128.

[99]. Davydov R, Sahlin M, Kuprin S, Gräslund A, Ehrenberg A (1996) Effect of the tyrosyl radical on the reduction and structure of the Escherichia coli ribonucleotide reductase protein R2 diferric site as probed by EPR on the mixed-valent state. Biochemistry 35, 5571–5576. [PubMed: 8611548]

[100]. Davydov RM, Davydov A, Ingemarson R, Thelander L, Ehrenberg A, Gräslund A (1997) EPR study of the mixed-valent diiron sites in mouse and herpes simplex virus ribonucleotide reductases. Effect of the tyrosyl radical on structure and reactivity of the diferric center. Biochemistry 36, 9093–9100. [PubMed: 9230041]

[101]. Davydov RM, Menage S, Fontecave M, Gräslund A, Ehrenberg A (1997) Mixed-valent μ-oxo-bridged diiron complexes produced by radiolytic reduction at 77 K studied by EPR. J. Biol. Inorg. Chem 2, 242–255.

[102]. Atta M, Andersson KK, Ingemarson R, Thelander L, Gräslund A (1994) EPR studies of mixed-valent Fe(III)-Fe(II) clusters formed in the R2 subunit of ribonucleotide reductase from mouse or herpes-simplex virus - mild chemical reduction of the diferric centers. J. Am. Chem. Soc 116, 6429–6430.

[103]. Horowitz S, Trievel RC (2012) Carbon-oxygen hydrogen bonding in biological structure and function. J. Biol. Chem 287, 41576–41582. [PubMed: 23048026]

[104]. Adhikari U, Scheiner S (2013) Magnitude and mechanism of charge enhancement of CH..O hydrogen bonds. J. Phys. Chem. A 117, 10551–10562. [PubMed: 24028630]

[105]. Bodea S, Funk MA, Balskus EP, Drennan CL (2016) Molecular basis of C-N bond cleavage by the glycyl radical enzyme choline trimethylamine-lyase. Cell Chem. Biol 23, 1206–1216. [PubMed: 27642068]

[106]. Horowitz S, Dirk LM, Yesselman JD, Nimtz JS, Adhikari U, Mehl RA, Scheiner S, Houtz RL, Al-Hashimi HM, Trievel RC (2013) Conservation and functional importance of carbon-oxygen hydrogen bonding in AdoMet-dependent methyltransferases. J. Am. Chem. Soc 135, 15536–15548. [PubMed: 24093804]

[107]. Starbird CA, Perry NA, Chen Q, Berndt S, Yamakawa I, Loukachevitch LV, Limbrick EM, Bachmann BO, Iverson TM, McCulloch KM (2018) The structure of the bifunctional everninomicin biosynthetic enzyme EvdMO1 suggests independent activity of the fused methyltransferase-oxidase domains. Biochemistry 57, 6827–6837. [PubMed: 30525509]

[108]. Marcotte EM, Pellegrini M, H.-L., N., Rice DW, Yeates TO, Eisenberg D (1999) Detecting protein function and protein-protein interactions from genome sequences. Science 285, 751–753. [PubMed: 10427000]

[109]. Maddocks SE, Oyston PC (2008) Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology 154, 3609–3623. [PubMed: 19047729]

[110]. Kulakova AN, Kulakov LA, McGrath JW, Quinn JP (2009) The construction of a whole-cell biosensor for phosphonoacetate, based on the LysR-like transcriptional regulator PhnR from *Pseudomonas fluorescens* 23F. Microb. Biotechnol 2, 234–240. [PubMed: 21261917]

[111]. Kulakova AN, Kulakov LA, Villarreal-Chiu JF, Gilbert JA, McGrath JW, Quinn JP (2009) Expression of the phosphonoalanine-degradative gene cluster from Variovorax sp Pal2 is induced by growth on phosphonoalanine and phosphonopyruvate. FEMS Microbiol. Lett 292, 100–106. [PubMed: 19191873]

[112]. Kulakova AN, Wisdom GB, Kulakov LA, Quinn JP (2003) The purification and characterization of phosphonopyruvate hydrolase, a novel carbon-phosphorus bond cleavage enzyme from *Variovorax sp* Pal2. J. Biol. Chem 278, 23426–23431. [PubMed: 12697754]

[113]. Mukhamedova KS, Glushenkova AI (2000) Natural phosphonolipids. Chem. Nat. Compd 36, 329–341.

[114]. Kirkpatrick DS, Bishop SH (1973) Phosphonoprotein - Characterization of aminophosphonic acid rich glycoproteins from sea anemones. Biochemistry 12, 2829–2840. [PubMed: 4146291]

[115]. Hilderbrand RL, Joseph JL, Lubansky HJ, Henderson TO (1977) Biology of Alkylphosphonic Acids, Naval Health Research Center, San Diego, CA.

[116]. Kittredge JS, Isbell AF, Hughes RR (1967) Isolation and characterization of N-methyl derivatives of 2-aminoethylphosphonic acid from sea anemone *Anthopleura xanthogrammica*. Biochemistry 6, 289–295. [PubMed: 4382153]

[117]. Kittredge JS, Roberts E (1969) A carbon-phosphorus bond in nature. Science 164, 37–42. [PubMed: 4388111]

[118]. Kataoka H, Sakiyama N, Makita M (1989) Determination of 2-aminoethylphosphonic acid and its N-methyl derivative in animal-tissues by gas-chromatography with flame photometric detection. Agr. Biol. Chem 53, 2791–2796.

[119]. McGrath JW, Chin JP, Quinn JP (2013) Organophosphonates revealed: new insights into the microbial metabolism of ancient molecules. Nat. Rev. Microbiol 11, 412–419. [PubMed: 23624813]

[120]. Quinn JP, Kulakova AN, Cooley NA, McGrath JW (2007) New ways to break an old bond: the bacterial carbon-phosphorus hydrolases and their role in biogeochemical phosphorus cycling. Environ. Microbiol 9, 2392–2400. [PubMed: 17803765]

[121]. Villarreal-Chiu JF, Quinn JP, McGrath JW (2012) The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment. Front Microbiol 3, 19. [PubMed: 22303297]

[122]. Peck SC, van der Donk WA (2013) Phosphonate biosynthesis and catabolism: a treasure trove of unusual enzymology. Curr. Opin. Chem. Biol 17, 580–588. [PubMed: 23870698]

[123]. Kononova SV, Nesmeyanova MA (2002) Phosphonates and their degradation by microorganisms. Biochemistry-Moscow 67, 184–195. [PubMed: 11952414]

[124]. Mus F, Crook MB, Garcia K, Garcia Costas A, Geddes BA, Kouri ED, Paramasivan P, Ryu MH, Oldroyd GE, Poole PS, Udvardi MK, Voigt CA, Ane JM, Peters JW (2016) Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. Appl. Environ. Microbiol 82, 3698–3710. [PubMed: 27084023]

[125]. Bardin S, Dan S, Osteras M, Finan TM (1996) A phosphate transport system is required for symbiotic nitrogen fixation by *Rhizobium meliloti*. J. Bacteriol 178, 4540–4547. [PubMed: 8755882]

[126]. Eckardt NA (2005) Insights into plant cellular mechanisms: Of phosphate transporters and arbuscular mycorrhizal infection. Plant Cell 17, 3213–3216.

[127]. Tang C, Hinsinger P, Drevon JJ, Jaillard B (2001) Phosphorus deficiency impairs early nodule functioning and enhances proton release in roots of *Medicago truncatula* L. Ann. Bot 88, 131–138.

[128]. Høgh-Jensen H, Schjoerring JK, Soussana JF (2002) The influence of phosphorus deficiency on growth and nitrogen fixation of white clover plants. Ann. Bot 90, 745–753. [PubMed: 12451030]

[129]. King GM (1988) Distributions and metabolism of quaternary amines in marine sediments, Blackburn TH, and Sørensen J, Ed., John Wiley & Sons, New York, NY, USA.

[130]. Bernard T, Pocard JA, Perroud B, Lerudulier D (1986) Variations in the response of salt-stressed *Rhizobium* strains to betaines. Arch. Microbiol 143, 359–364.

[131]. Lerudulier D, Bernard T (1986) Salt tolerance in Rhizobium - a possible role for betaines. FEMS Microbiol. Lett 39, 67–72.

[132]. Lerudulier D, Bernard T, Goas G, Hamelin J (1984) Osmoregulation in Klebsiella pneumoniae - Enhancement of anaerobic growth and nitrogen-fixation under stress by proline betaine, gamma-butyrobetaine, and other related compounds. Can. J. Microbiol 30, 299–305. [PubMed: 6372974]

[133]. Boncompagni E, Osteras M, Poggi MC, Le Rudulier D (1999) Occurrence of choline and glycine betaine uptake and metabolism in the family Rhizobiaceae and their roles in osmoprotection. Appl. Environ. Microbiol 65, 2072–2077. [PubMed: 10224003]

[134]. Goldmann A, Boivin C, Fleury V, Message B, Lecoeur L, Maille M, Tepfer D (1991) Betaine use by rhizosphere bacteria - Genes essential for trigonelline, stachydrine, and carnitine catabolism in *Rhizobium meliloti* are located on Psym in the symbiotic Region. Mol. Plant Microbe 4, 571–578.

[135]. Ticak T, Kountz DJ, Girosky KE, Krzycki JA, Ferguson DJ (2014) A nonpyrrolysine member of the widely distributed trimethylamine methyltransferase family is a glycine betaine methyltransferase. Proc. Natl. Acad. Sci. U S A 111, E4668–4676. [PubMed: 25313086]

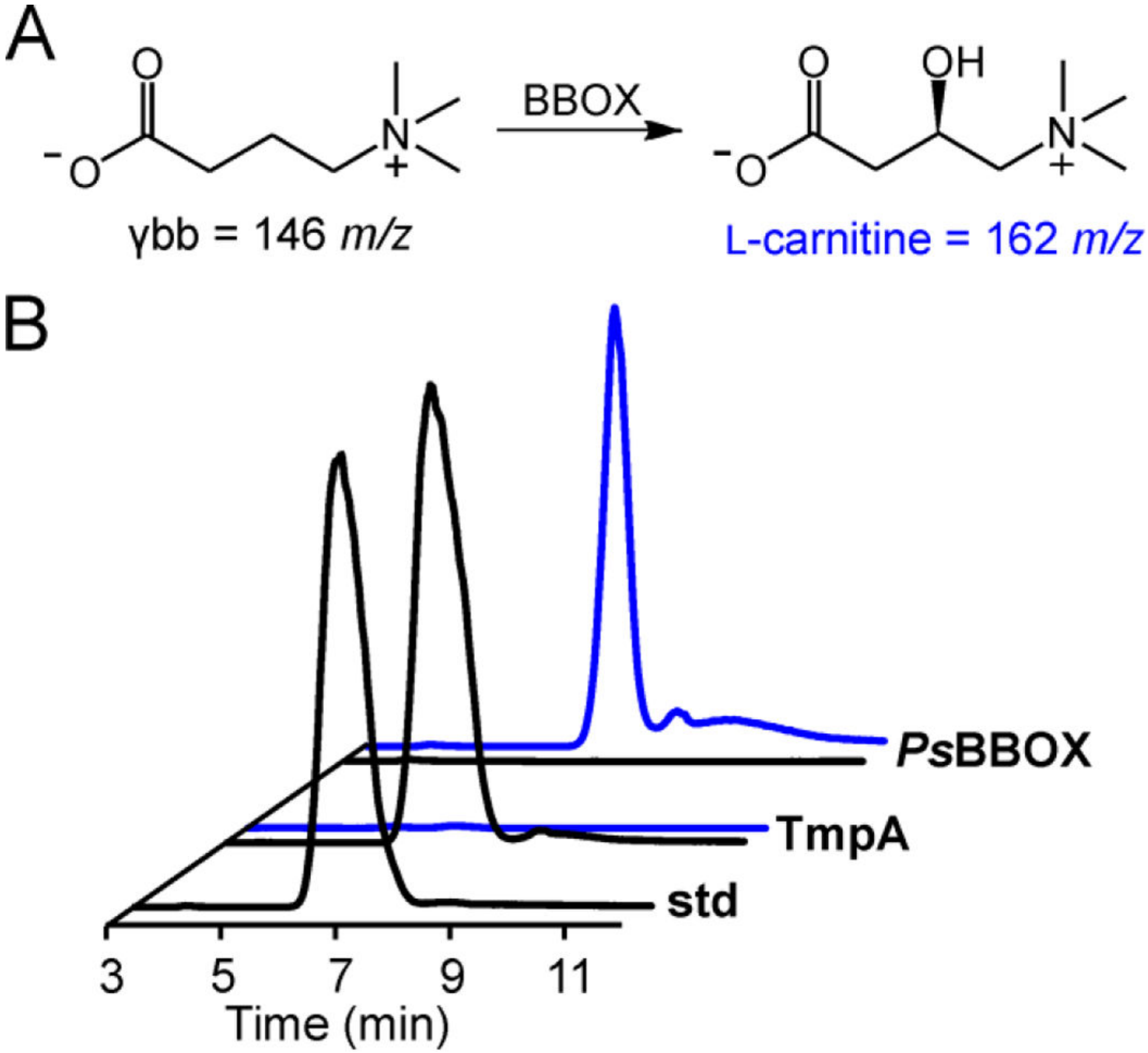


Figure 1.

Assessment of the γbb hydroxylase activities of TmpA and *Ps* BBOX. (A) The reaction catalyzed by BBOX. (B) LC-MS chromatograms monitoring γbb (146 *m/z*, black) and L-carnitine (162 *m/z*, blue) after a 4-h incubation of 0.01 mM TmpA or *Ps* BBOX, 0.02 mM $(NH_4)_2Fe(SO_4)_2$, 0.2 mM sodium L-ascorbate, 1 mM 2OG and 1 mM γbb.

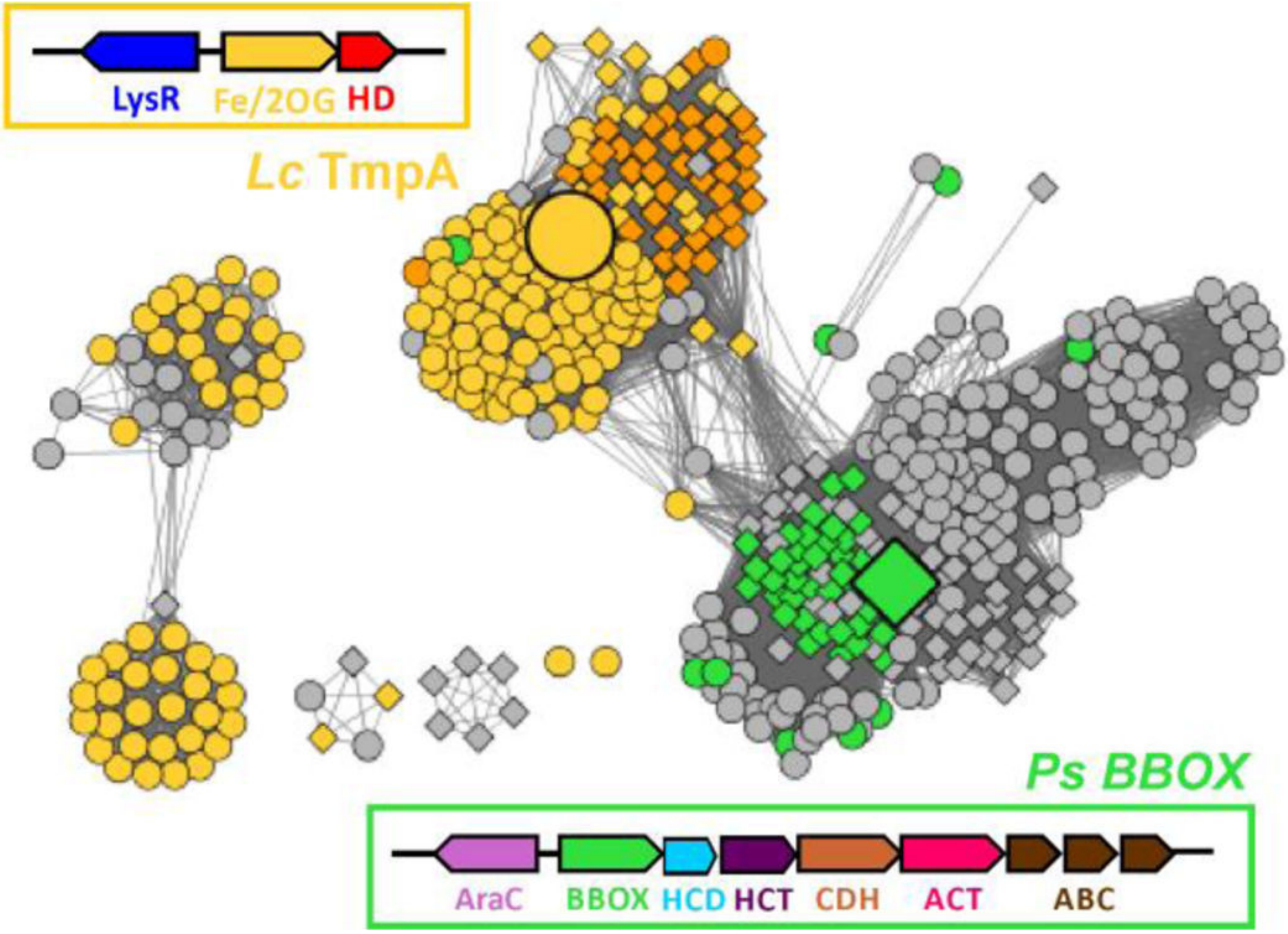


Figure 2.

Sequence similarity network (SSN) illustrating divergence of BBOX-like and TmpA-like proteins. The clusters of bacterial protein sequences shown here are derived from the IPR003819 SSN (Figures S1–3). The nodes represent protein sequences with > 90% identity. Edges between the nodes represent pairwise alignment scores of $< 10^{-71}$ (corresponding to ~ 33% sequence identity). The large yellow circle and green diamond represent *Lc* TmpA and *Ps* BBOX, respectively. Orange nodes represent predicted fusion proteins with both TmpA- and TmpB-like domains. Diamond shaped nodes represent sequences with an N-terminal Zn(II)-binding motif (see discussion). The yellow and green boxes summarize the genomic context of proteins represented by the yellow and green nodes, respectively. Other designations: LysR, LysR-type transcription regulator; AraC, AraC-type transcription regulator; CDH, carnitine dehydrogenase; HCT, 4-hydroxybenzoyl-CoA thioesterase; HCD, 3-hydroxybutyryl-CoA dehydrogenase; ACT, Acetyl-CoA acetyltransferase/thiolase; ABC, ABC-type glycine/betaine transporter.

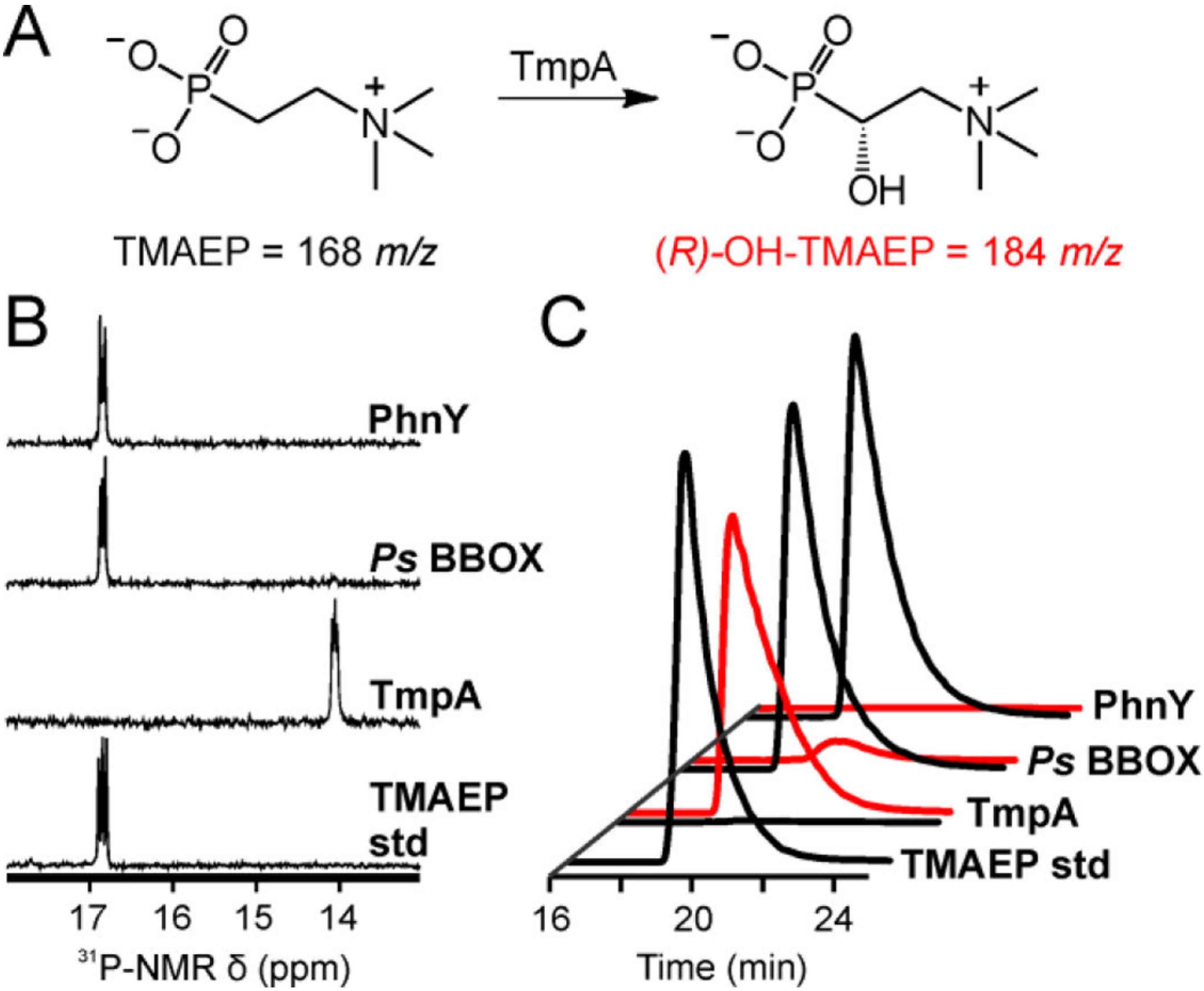


Figure 3.

Activities of selected Fe/2OG oxygenases toward TMAEP. (A) The chemical transformation catalyzed by TmpA. (B) ^{31}P-NMR spectra and (C) LC-MS chromatograms after a 4-h incubation of a solution containing 0.01 mM TmpA, PhnY or *Ps* BBOX with 0.02 mM $(NH_4)_2Fe(SO_4)_2$, 0.2 mM L-ascorbate, 3 mM 2OG and 2 mM TMAEP. The triplet splitting in the ^{1}H-decoupled ^{31}P-NMR spectra of the TMAEP substrate and (*R*)-OH-TMAEP product in panel B is presumed to result from the ^{14}N nucleus on C1. The reason it is present in the spectrum of the *N*-trimethylated compound, but not those of the di- and unmethylated analogs, is not clear and is discussed more in the Supporting Information.

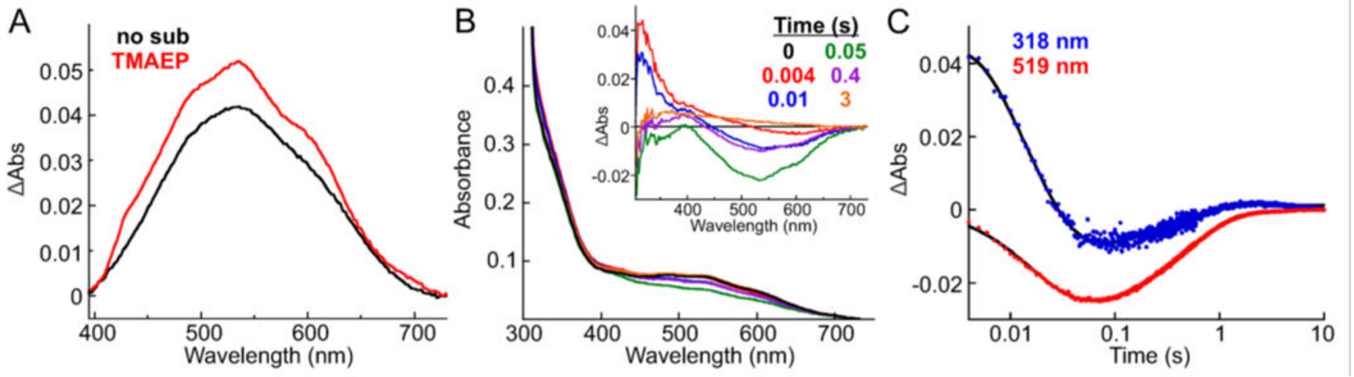


Figure 4.

Ultraviolet-visible absorption data showing binding of TMAEP to TmpA•Fe(II)•2OG and triggering of O_2 addition. (A) Difference absorption spectra associated with binding of 2OG (5 mM) to the TmpA•Fe(II) complex [0.6 mM TmpA, 0.5 mM $(NH_4)_2Fe(SO_4)_2$] in the absence of substrate (black) and presence of 5 mM TMAEP (red). (B) Stopped-flow absorption spectra acquired after mixing at 5 °C of a solution of 1.2 mM TmpA, 1 mM $(NH_4)_2Fe(SO_4)_2$, 10 mM 2OG and 10 mM TMAEP with an equal volume of air-saturated 50 mM sodium HEPES buffer, pH 7.5 (~ 0.4 mM O_2 at 5 °C). The inset shows the absorption spectra at indicated time points after subtraction of the spectrum of the TmpA•Fe(II)•2OG•TMAEP complex. (C) Kinetic traces at 318 nm (blue dots) and 519 nm (red dots) extracted from the time-dependent spectra in panel B. The solid black lines are regression fits to the data, as described in the Experimental Section.

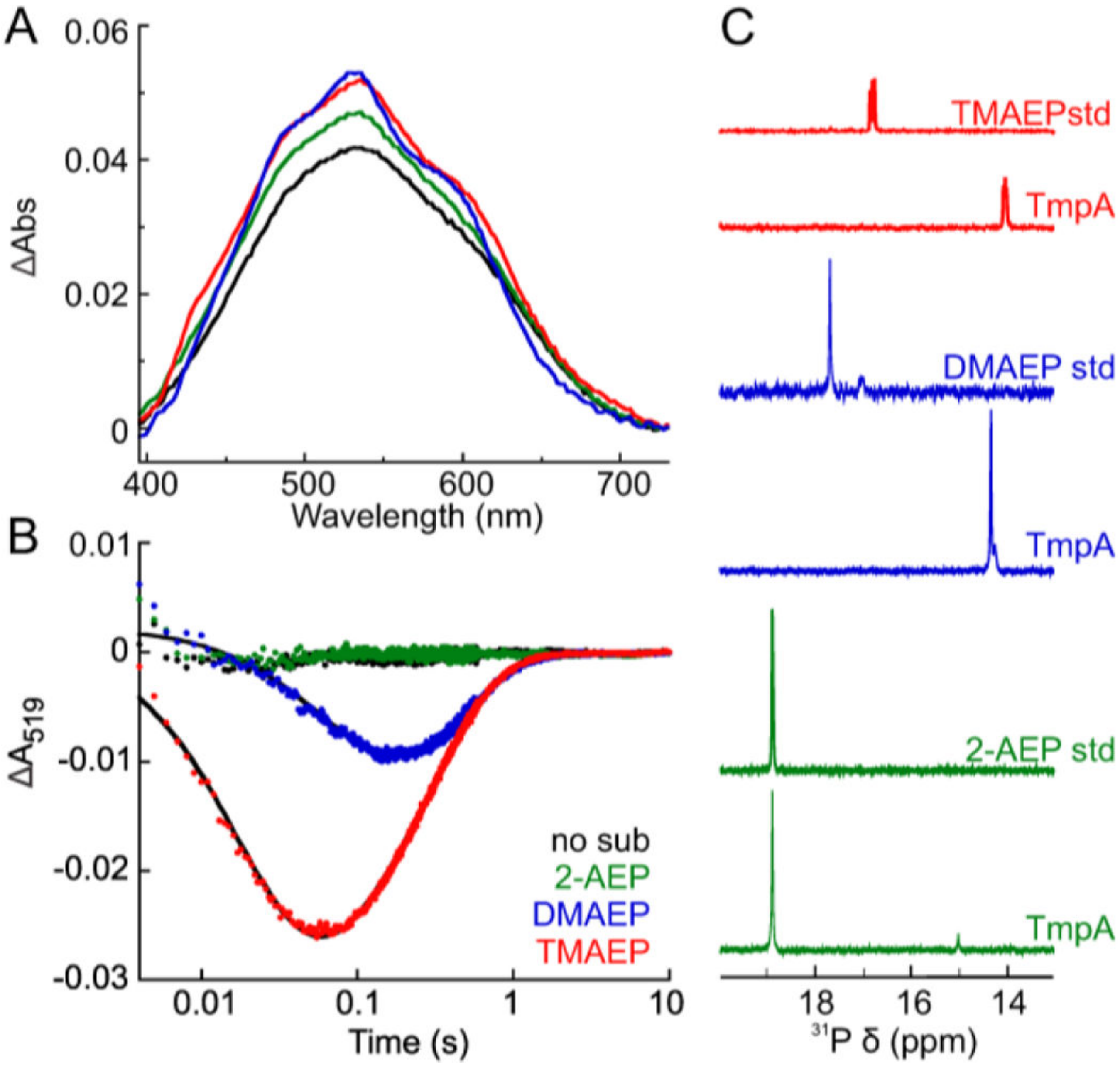


Figure 5.

Activity of TmpA toward TMAEP analogues with varying degrees of *N*-methylation. (A) Absorption difference spectra caused by binding of 2OG (5 mM) to the TmpA•Fe(II) complex [0.6 mM TmpA, 0.5 mM $(NH_4)_2Fe(SO_4)_2$] in the absence of a substrate (black) and in the presence of 5 mM TMAEP (red), DMAEP (blue) or 2-AEP (green). (B) Kinetic traces at 519 nm after mixing of the solutions described in panel A with air-saturated 50 mM sodium HEPES buffer, pH 7.5 (~ 0.2 mM O_2 final at 5 °C). The solid black lines are non-linear regression fits to the data, as described in the Experimental Section. (C) ^{31}P-NMR spectra of reaction samples after a 4 h incubation of a solution of 0.01 mM TmpA, 0.02 mM

$(NH_4)_2Fe(SO_4)_2$, 0.2 mM sodium ascorbate, 3 mM 2OG and 2 mM of either TMAEP (red), DMAEP (blue), or 2-AEP (green).

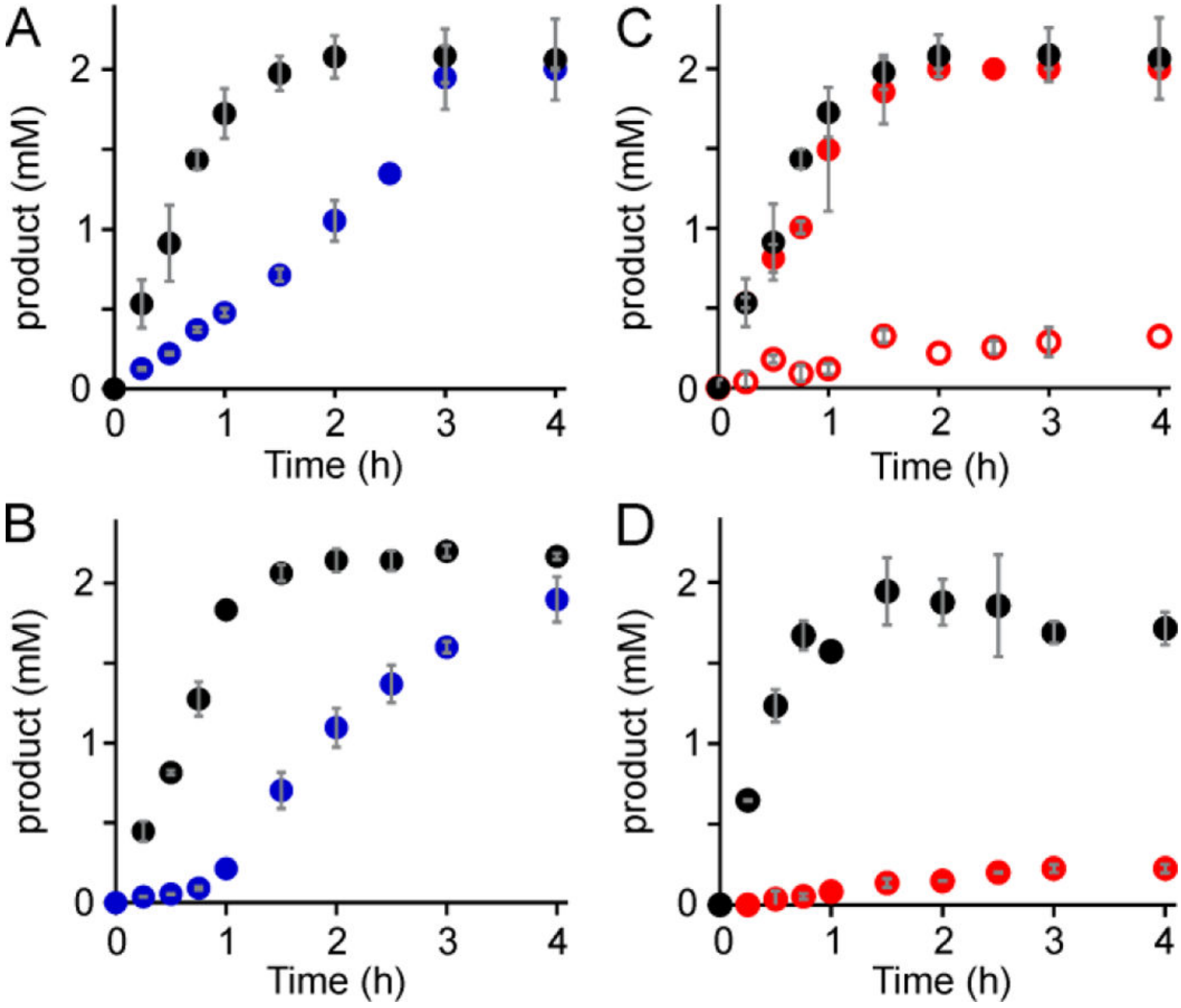


Figure 6.

Evaluation of relative specificity of TmpA for TMAEP and either DMAEP or PC by direct competition. All reactions were performed at 3 °C and contained 0.02 mM TmpA, 0.03 mM $(NH_4)_2Fe(SO_4)_2$, 0.4 mM sodium ascorbate, 6 mM 2OG and 2 mM of each substrate. (A) Control reactions containing either TMAEP (black) or DMAEP (blue) in the absence of the other compound, monitoring corresponding hydroxylated products by LC-MS. (B) Competition reactions containing both TMAEP and DMAEP, monitoring corresponding hydroxylated products by LC-MS. (C) Control reactions containing either TMAEP (black) or PC (filled red), monitoring (*R*)-OH-TMAEP production and PC consumption, respectively, by LC-MS The open red circles are the concentration of PC consumed, detected by LC-MS, from a reaction containing PC and 2 mM (*R*)-OH-TMAEP. (D) Competition reactions containing both TMAEP and PC, monitoring (*R*)-OH-TMAEP and phosphate products by ^{31}P-NMR because PC and (*R*)-OH-TMAEP have the same *m/z* and similar retention times.

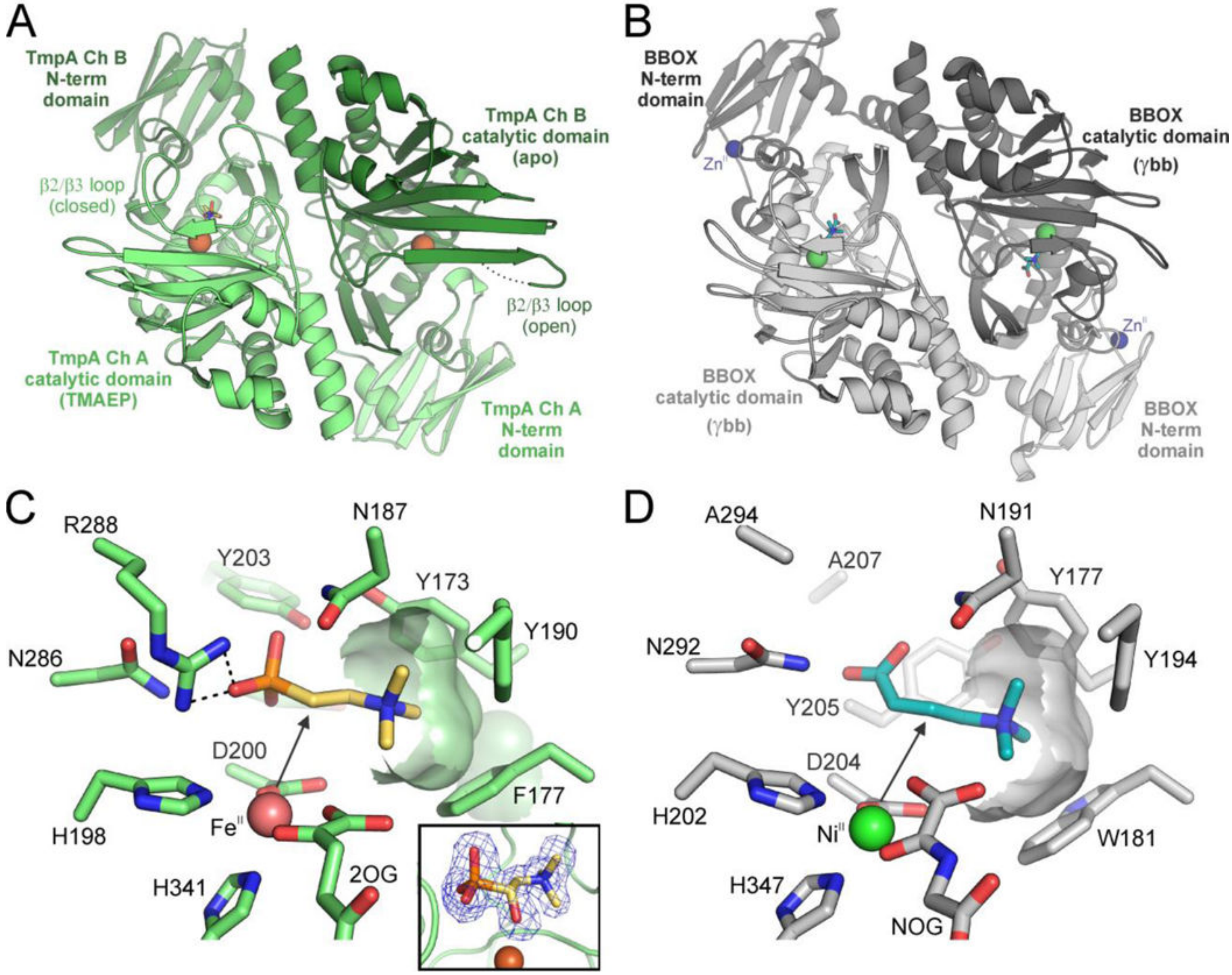


Figure 7.

Structural comparison of *Lc* TmpA and *Hs* BBOX. (A) Homodimeric quaternary structure of TmpA with chain A in light green and chain B in dark green. Fe(II) ions are shown as brown spheres. (B) Homodimeric quaternary structure of *Hs* BBOX (PDB accession code 3MS5) with chain A in dark gray and chain B in light gray. The Zn(II) and Ni(II) ions are shown as blue and green spheres, respectively. (C) Active site of chain A in the TmpA substrate-bound structure, showing the co-substrate (2OG), amino acid side chains, and TMAEP (yellow) in stick format. Electrostatic interactions are designated by black dashed lines and the black arrow identifies the position of hydroxylation. The inset shows the F_o-F_c omit map contoured at 3.0σ (blue mesh) and atomic model for (*R*)-OH-TMAEP (yellow sticks) in chain A of the TmpA product-bound structure. (D) Active site in the substrate-bound structure of *Hs* BBOX (PDB accession code 3MS5), showing the co-substrate analog, *N*-oxalylglycine (NOG), amino acids side chains, and γbb (blue) in stick format.

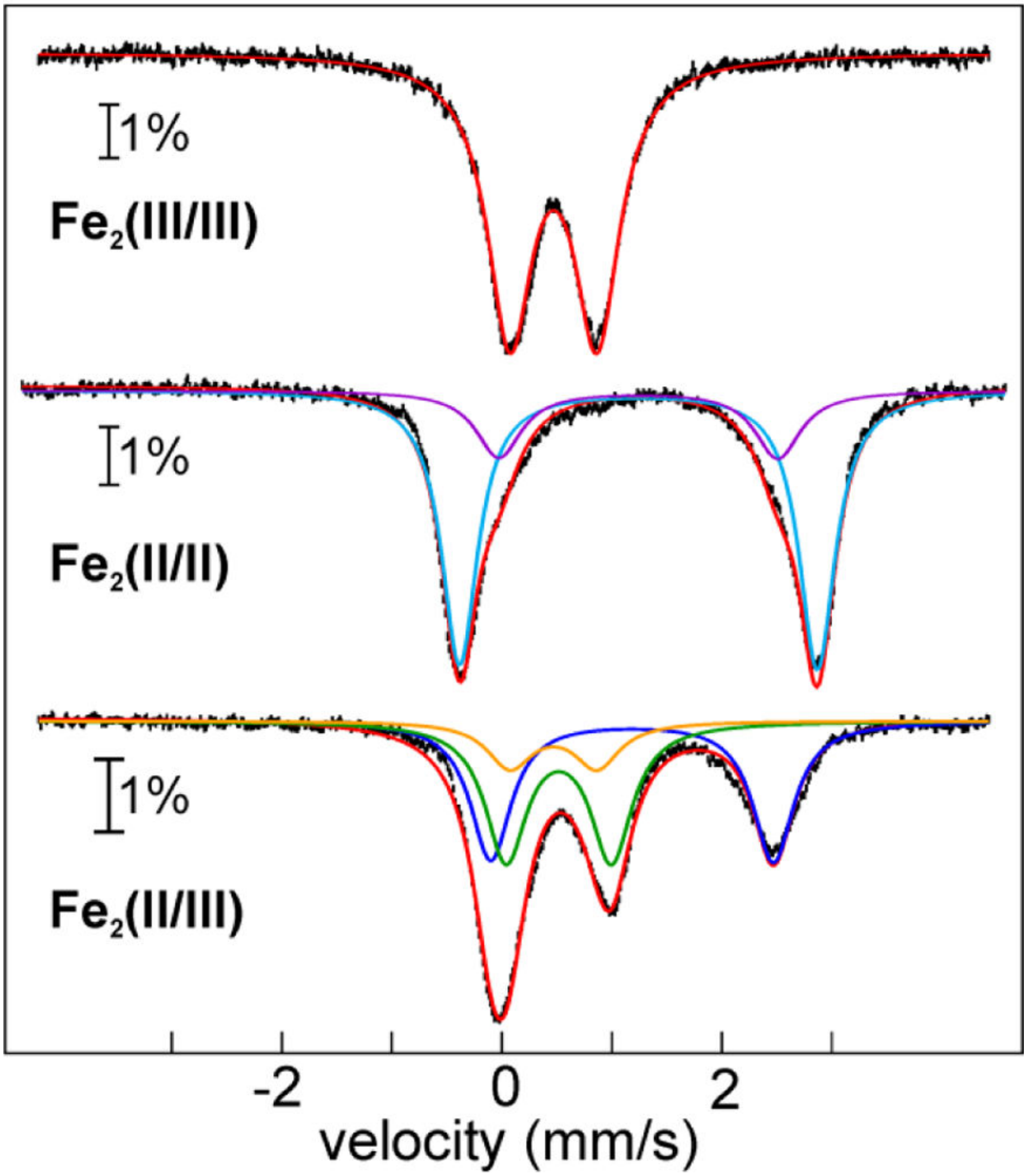


Figure 8.

120-K/0-T Mössbauer spectra of 2 mM O_2-free TmpA prepared in three oxidation states. All incubations were carried out for 45 min in an anoxic chamber. Experimental spectra are shown as black vertical bars. Overall simulations are shown as red lines. (*Top*) Fe_2(III/III) state obtained by treatment of the as-isolated protein with 3 mM potassium ferricyanide. (*Middle*) Fe_2(II/II) state obtained by treatment of the as-isolated protein with 20 mM sodium dithionite. The two quadrupole-doublet components of the simulated spectrum of the Fe_2(II/II) complex are shown as purple and cyan lines; the parameters are provided in the main text. (*Bottom*) Fe_2(II/III) state accumulated to ~ 65 % by treatment of the as-isolated protein with 20 mM sodium L-ascorbate. The simulated sub-spectra corresponding to the Fe_2(II/III) species (orange) and the paired Fe(II) (blue) and Fe(III) (green) sub-sites are shown as solid lines; their parameters are provided in the main text. The contribution of the Fe_2(II/II) species (22%) has already been removed (by subtraction from the experimental spectrum; Figure S20) for clarity.

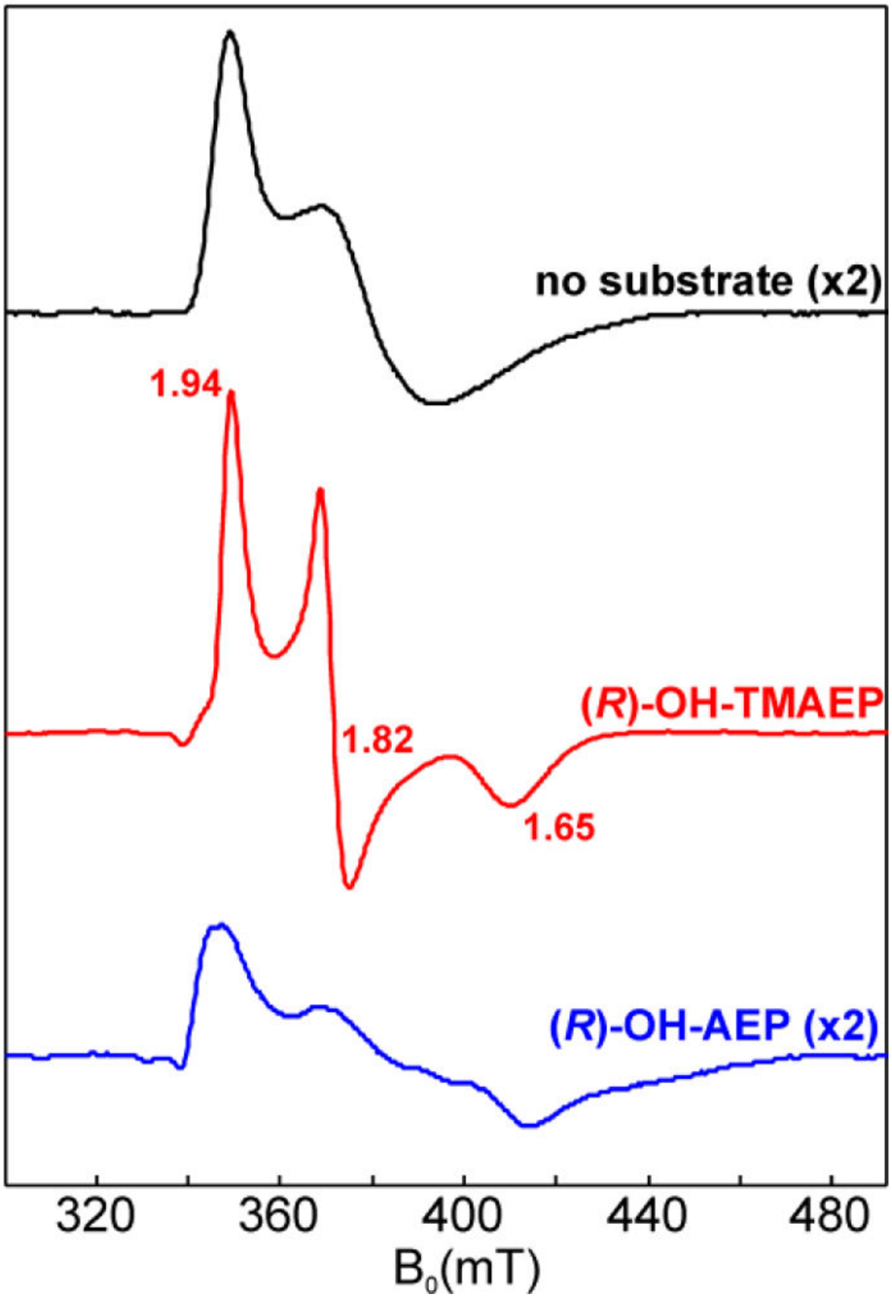


Figure 9.

X-band EPR spectra of O_2-free 0.25 mM TmpB after treatment with 10 mM sodium L-ascorbate (black) followed by addition of either 10 mM (*R*)-OH-TMAEP (red) or (*R*)-OH-AEP (blue). Experimental conditions: temperature = 10 K, microwave power = 0.2 mW, microwave frequency = 9.479 GHz, modulation amplitude = 1 mT.

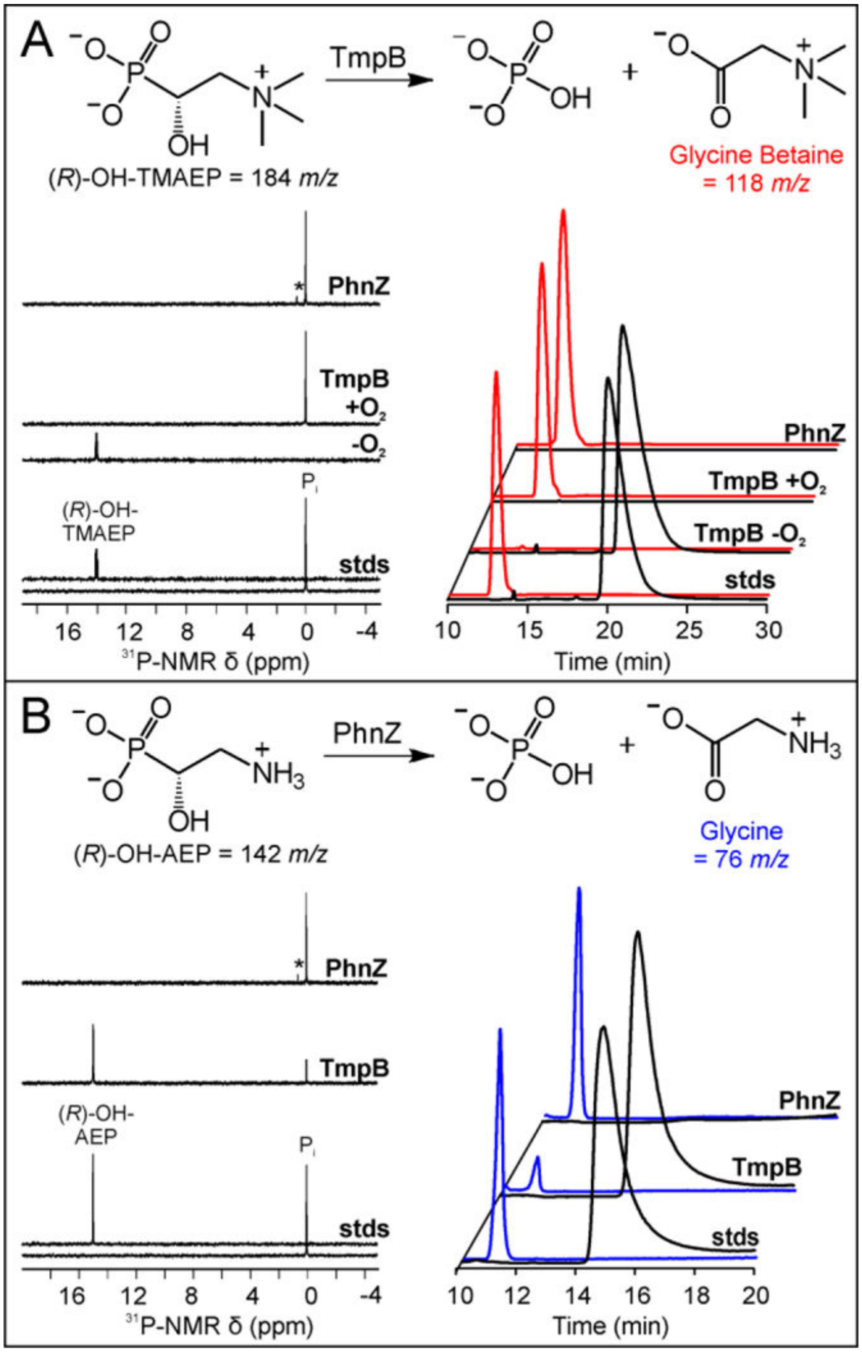


Figure 10.

Activity of the HD-MVDOs against aminophosphonate compounds with and without *N*-methylation. The chemical transformations are depicted at the top of each panel. Aerobic reactions containing 0.01 mM TmpB or PhnZ, 0.2 mM L-ascorbate, and 2 mM of either (A) (*R*)-OH-TMAEP or (B) (*R*)-OH-AEP were incubated for 4 h at 3 °C. (*Left*) ^{31}P-NMR spectra detecting the substrates, (*R*)-OH-TMAEP or (*R*)-OH-AEP, and the product phosphate. The asterisks mark a contaminant present in the PhnZ protein preparation.

(*Right*) LC-MS chromatograms detecting the substrates, (*R*)-OH-TMAEP or (*R*)-OH-AEP, and products, glycine betaine or glycine.

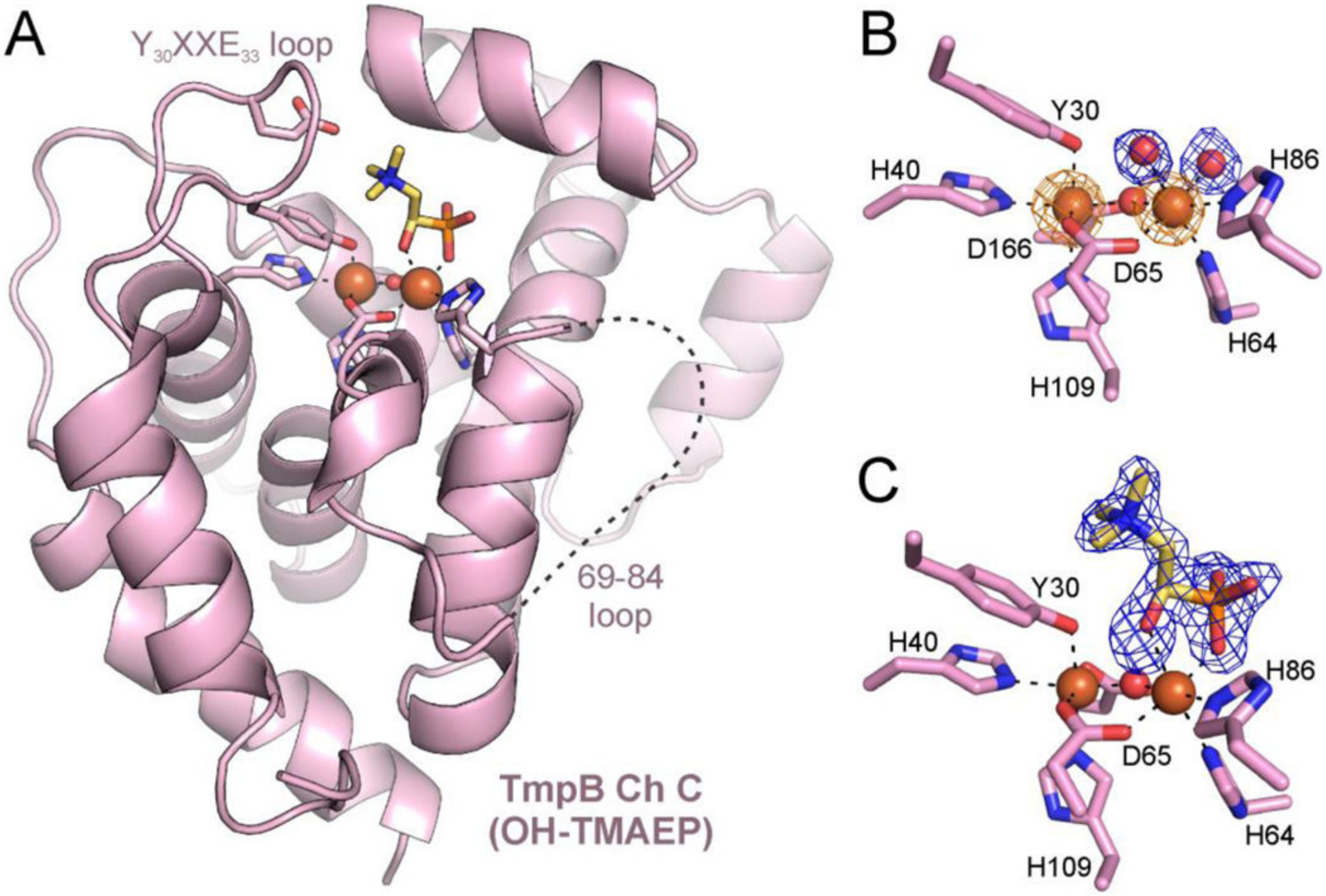
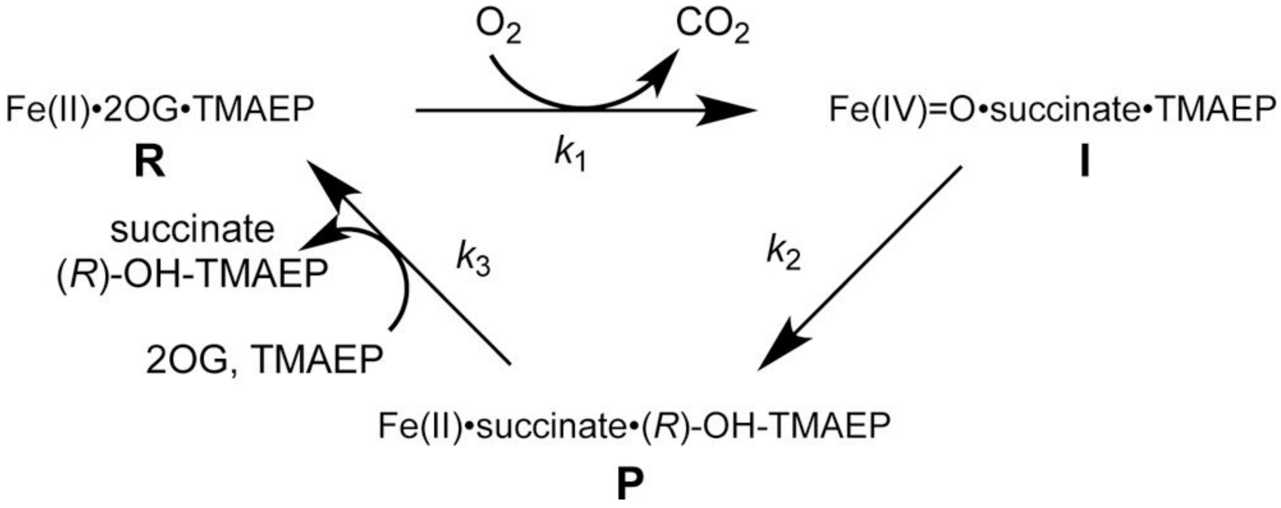


Figure 11.

Views of TmpB from the X-ray crystal structure solved in this study. (A) Cartoon depiction of chain C of the TmpB structure. Iron ions are shown as brown spheres and (*R*)-OH-TMAEP (yellow) is shown in stick format. A dashed line represents the unmodeled region of the structure. TmpB active site of (B) chain A lacking substrate and (C) chain D with (*R*)-OH-TMAEP (yellow sticks) bound. Water molecules are shown as red spheres and iron ions as orange spheres. Orange mesh depicts the anomalous Fourier density contoured at 10σ. Blue mesh depicts the $F_o - F_c$ omit map contoured at 3σ for either the coordinating water molecules or the substrate.


O2
CO2
Fe(II)•2OG•TMAEP
R
k1
Fe(IV)=O•succinate•TMAEP
I
k2
succinate
(R)-OH-TMAEP
k3
2OG, TMAEP
Fe(II)•succinate•(R)-OH-TMAEP
P

Scheme 1.