

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

**SUPPORTING INFORMATION**

*Lauren J. Rajakovich<sup>†#</sup>, Maria-Eirini Pandelia<sup>‡<sup>†</sup></sup>, Andrew J. Mitchell<sup>†<sup>†</sup></sup>, Wei-chen Chang<sup>‡§</sup>, Bo  
Zhang<sup>‡<sup>†</sup></sup>, Amie K. Boal<sup>†<sup>†</sup>\*</sup>, Carsten Krebs<sup>†<sup>†</sup>\*</sup>, J. Martin Bollinger, Jr.<sup>†<sup>†</sup>\*</sup>*

<sup>†</sup>Department of Biochemistry and Molecular Biology and <sup>‡</sup>Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, United States

<sup>#</sup>Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States

<sup>†<sup>†</sup></sup>Present address: Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02453, United States

<sup>†<sup>†</sup></sup>Present address: Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142

<sup>§</sup>Present address: Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695, United States

<sup>†<sup>†</sup></sup>Present address: REG Life Sciences, LLC, South San Francisco, California 94080

**\*Corresponding Authors**

akb20@psu.edu

cdk10@psu.edu

jmb21@psu.edu

## Table of Contents

Supplementary Experimental Procedures.....	S3-7
Preparation of <i>Lc</i> TmpA	
Preparation of <i>Lc</i> TmpB	
Preparation of <i>Ps</i> BBOX	
Synthesis of DMAEP and TMAEP	
Supplementary Tables.....	S8-10
Table S1: Target <i>m/z</i> ratios in LC-MS analysis	
Table S2: Target <i>m/z</i> ratios in LC-MS analysis	
Table S3: TmpA x-ray structure data collection and refinement statistics	
Table S4: TmpB x-ray structure data collection and refinement statistics	
Supplementary Figures.....	S11-35
Figures S1-3: Sequence Similarity Network analysis	
Figure S4: NMR characterization of OH-TMAEP	
Figure S5: TmpA activity with $\gamma$ bb analogues	
Figure S6: TmpA activity with phosphocholine	
Figure S7: TmpA reaction requirements for hydroxylation of TMAEP	
Figure S8: TmpA O <sub>2</sub> reactivity without substrate	
Figure S9-11: TmpA reactivity with TMAEP analogues	
Figure S12-17: TmpA x-ray structure analysis	
Figure S18: Comparison of TmpB and PhnZ primary structure	
Figure S19: Evidence for adventitiously-bound Fe(III) in TmpB	
Figure S20: Mössbauer-spectroscopic analysis of TmpB cofactor	
Figure S21: TmpB activity toward OH-DMAEP	
Figure S22-24: TmpB x-ray structure analysis	
Figure S25: TmpB absorption spectrum in the presence of substrate	
References.....	S36

**Preparation of *Lc* TmpA.** The *Lc* gene CAER\_RS0122985 encoding TmpA (NCBI RefSeq WP\_027237574.1) was codon-optimized for over-expression in *Escherichia coli* (*Ec*), synthesized, and inserted into the *NdeI* and *XhoI* restriction sites of expression vector pET-28a(+) by GenScript (Piscataway, NJ). *Ec* BL21 (DE3) competent cells (Invitrogen; Carlsbad, CA) were transformed with the TmpA plasmid and selected for kanamycin resistance. Transformed cells were grown in rich Luria-Bertani medium (LB) with 50 mg/L kanamycin at 37 °C with shaking (250 rpm) until an OD<sub>600</sub> of 0.6-0.8 was reached. Protein expression was then induced by addition of 0.25 mM IPTG, and cell cultures were incubated at 18 °C with shaking (250 rpm) for 16-20 h. Cultures were centrifuged at 8,000 × *g* for 15 min; cell pellets were flash frozen in liquid N<sub>2</sub> and stored at -80 °C. Cell pellets were resuspended in 50 mM Tris-HCl (pH 7.5) buffer containing 150 mM NaCl and 10% glycerol (buffer A). The cells were lysed by passage of the suspension through a Microfluidics M-110EH-30 microfluidizer processor at 20,000 psi for 10 min and centrifuged at 22,000 × *g* for 20 min. The supernatant was loaded onto a Ni(II)-NTA immobilized affinity chromatography column (~100 mL resin per 500 mL lysate) in buffer A. The column was washed with 3 column volumes of buffer A containing 30 mM imidazole. The protein was eluted with buffer A containing 250 mM imidazole. Fractions containing the protein were pooled and concentrated at 3,500 × *g* using a 30K MWCO Macrosep<sup>®</sup> Advance Centrifugal Device (Pall Corporation, Port Washington, NY). The concentrated protein was then dialyzed against 100 equivalent volumes of buffer A containing 5 mM EDTA and 10 mM dithiothreitol (DTT) and then exchanged back into buffer A by two additional rounds of dialysis. Protein for crystallography was further purified by size exclusion chromatography using a Sephadex S-200 column (120 mL resin) equilibrated in buffer A lacking glycerol. Fractions containing the protein were pooled and concentrated at 3,500 × *g* using a 30K MWCO Macrosep<sup>®</sup> Advance Centrifugal Device. The

protein was frozen in liquid N<sub>2</sub> and stored at -80 °C. Protein purity was assessed by SDS-PAGE with Coomassie staining, and protein concentration was determined by using a molar absorption coefficient of 49,500 M<sup>-1</sup>•cm<sup>-1</sup> at 280 nm.<sup>1</sup> The iron content was determined by the ferrozine assay<sup>2</sup> to be < 0.08 Fe/protein and considered to be “apo”.

**Preparation of *Lc* TmpB.** The *Lc* gene CAER\_RS0122980 encoding TmpB (NCBI RefSeq WP\_027237573.1) was codon-optimized for over-expression in *Ec*, synthesized, and inserted into the *NdeI* and *XhoI* restriction sites of expression vector pET-28a(+) by GenScript. *Ec* BL21 (DE3) competent cells were transformed with the TmpB plasmid and selected for kanamycin resistance. Transformed cells were cultured at 37 °C with shaking (250 rpm) in M9 minimal medium with 50 mg/L kanamycin, 150 mg/L ampicillin, 0.2% (v/v) glucose, 0.10 mM CaCl<sub>2</sub>, 200 mM MgSO<sub>4</sub>·7H<sub>2</sub>O, and 0.125 mM (NH<sub>4</sub>)<sub>2</sub>Fe(SO<sub>4</sub>)<sub>2</sub>·6H<sub>2</sub>O until an OD<sub>600</sub> of 0.6-0.8 was reached. Protein expression was then induced by addition of 0.1 mM IPTG, and cultures were also supplemented with 0.125 mM (NH<sub>4</sub>)<sub>2</sub>Fe(SO<sub>4</sub>)<sub>2</sub>·6H<sub>2</sub>O at this time. The cultures were incubated at 18 °C for 24 h with shaking (250 rpm) and then were centrifuged at 8,000 × *g* for 15 min. Cell pellets were flash frozen in liquid N<sub>2</sub> and stored at -80 °C. Cell pellets were resuspended in 50 mM Tris-HCl (pH 7.5) buffer containing 150 mM NaCl and 10% glycerol. The cells were lysed by passage through a Microfluidics M-110EH-30 microfluidizer processor at 20,000 psi for 10 min and centrifuged at 22,000 × *g* for 20 min. The supernatant was loaded onto a Ni(II)-NTA immobilized affinity chromatography column (~100 mL resin per 500 mL lysate) equilibrated in buffer A. The column was washed with 3 column volumes of buffer A. The protein was then eluted with buffer A containing 250 mM imidazole. Fractions containing the protein were pooled and concentrated at 3,500 × *g* using a 10K MWCO Macrosep<sup>®</sup> Advance Centrifugal Device. The concentrated protein was then dialyzed three times against 100 equivalent volumes of buffer A for

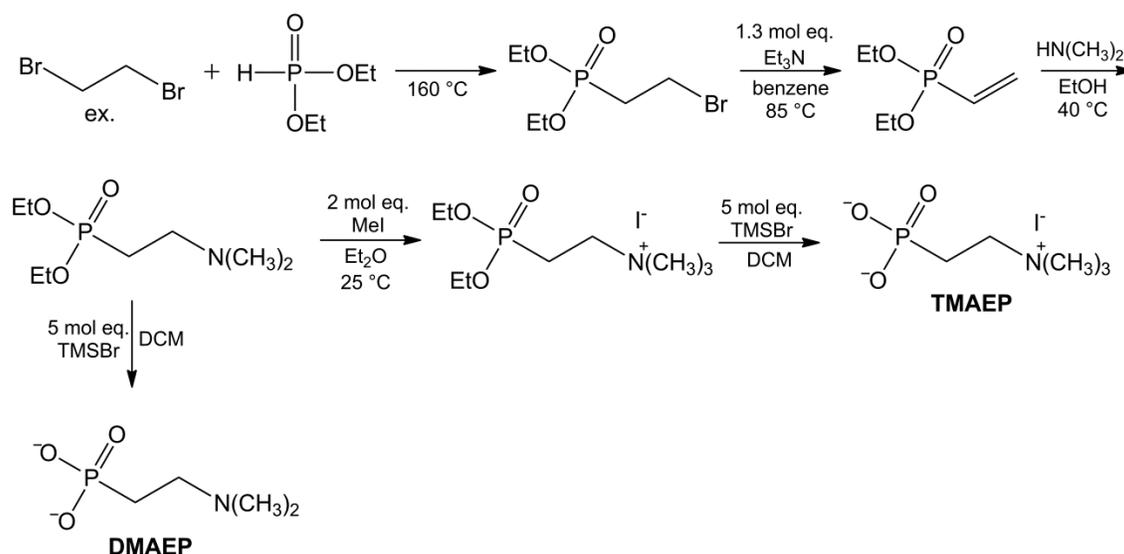
> 6 h. Protein for crystallography was further purified by size-exclusion chromatography using a 120 mL S-200 Sephadex column (GE) equilibrated in buffer A lacking glycerol. Fractions containing the protein were pooled and concentrated at  $3,500 \times g$  using a 10K MWCO Macrosep<sup>®</sup> Advance Centrifugal Device. The protein was frozen in liquid N<sub>2</sub> and stored at -80 °C. Protein purity was assessed by SDS-PAGE with Coomassie staining, and protein concentration was determined by using a molar absorption coefficient of  $14,650 \text{ M}^{-1} \cdot \text{cm}^{-1}$  at 280.<sup>1</sup> The iron content was determined to be 0.9-1.3 Fe/protein by ferrozine assay<sup>2</sup> and was confirmed by inductively coupled plasma atomic emission spectroscopy (ICP-AES), which detected less than 0.01 ppm of Mn, Zn and Cu. The concentrations of TmpB quoted in all experiments represent the concentration of diiron cluster, calculated as half the measured iron concentration.

**Preparation of *Ps* BBOX.** *Ec* BL21 (DE3) competent cells were transformed with the plasmid containing the *Ps* BBOX gene in a pCOLDi vector and selected for ampicillin resistance. Transformed cells were grown in rich Luria-Bertani medium (LB) with 150 mg/L ampicillin at 37 °C with shaking (250 rpm) until an OD<sub>600</sub> of 0.6-0.8 was reached. Protein expression was then induced by addition of 0.25 mM IPTG, and cell cultures were incubated at 18 °C with shaking (250 rpm) for 16-20 h. Cultures were centrifuged at  $8,000 \times g$  for 15 min; cell pellets were flash frozen in liquid N<sub>2</sub> and stored at -80 °C. Cell pellets were resuspended in 50 mM Tris-HCl (pH 7.5) buffer containing 150 mM NaCl and 10% glycerol. The cells were lysed by passage through a French press cell twice at >12,000 psi and centrifuged at  $22,000 \times g$  for 20 min. The supernatant was loaded onto a Ni(II)-NTA immobilized affinity chromatography column (~100 mL resin per 500 mL lysate) equilibrated in buffer A. The column was washed with 3 column volumes of buffer A containing 30 mM imidazole. The protein was then eluted with buffer A containing 250 mM imidazole. Fractions containing the protein were pooled and concentrated at  $3,500 \times g$  using a 30K

MWCO Macrosep<sup>®</sup> Advance Centrifugal Device. The protein was buffer exchanged to remove imidazole using a PD10 desalting column equilibrated in buffer A lacking glycerol. The protein was further purified by size-exclusion chromatography using a Sephadex S-200 column (120 mL resin) equilibrated in buffer A lacking glycerol. Fractions containing the protein were pooled and concentrated at  $3,500 \times g$  using a 30K MWCO Macrosep<sup>®</sup> Advance Centrifugal Device. The protein was frozen in liquid N<sub>2</sub> and stored at -80 °C. Protein purity was assessed by SDS-PAGE with Coomassie staining, and protein concentration was determined by using a molar absorption coefficient of  $55,190 \text{ M}^{-1} \cdot \text{cm}^{-1}$  at 280 nm.<sup>1</sup> Iron content was determined to be  $< 0.02 \text{ Fe/protein}$  by the ferrozine assay<sup>2</sup> and to be  $< 0.01$  by inductively coupled plasma atomic emission spectroscopy (ICP-AES). Zinc content was determined to be  $0.89 \text{ Zn/protein}$  by ICP-AES.

**Synthesis of 2-(dimethylamino)ethylphosphonate (DMAEP) and 2-(trimethylammonio)-ethylphosphonate (TMAEP).** DMAEP and TMAEP were synthesized by procedures adapted from previously published approaches (Scheme S1).<sup>3-6</sup> Triethylphosphite was reacted with excess dibromoethane at 160 °C under reflux for 6 h. The diethyl 2-bromoethylphosphonate product [<sup>31</sup>P NMR (CDCl<sub>3</sub>)  $\delta = 25.6 \text{ ppm}$ ] was isolated from the reaction mixture by reduced-pressure distillation. The product was mixed with 1.3 molar equiv. of triethylamine in benzene and refluxed at 85 °C for 2 h to give diethyl vinylphosphonate [<sup>31</sup>P NMR (CDCl<sub>3</sub>)  $\delta = 17.3 \text{ ppm}$ ]. The reaction was filtered, and the solid was washed and dried using a rotary evaporator. The product was reacted with 2 molar equiv. of dimethyl amine in ethanol while refluxing at 40 °C overnight to give diethyl 2-(dimethylamino)ethylphosphonate [<sup>31</sup>P NMR (CDCl<sub>3</sub>)  $\delta = 30.6 \text{ ppm}$ ]. The reaction was dried using a rotary evaporator, and the product was either directly deprotected (see below) to give DMAEP [<sup>31</sup>P NMR (CDCl<sub>3</sub>)  $\delta = 17.5 \text{ ppm}$ ] or incubated with 2 molar equiv. of methyl iodide in diethyl ether at 25 °C overnight to give diethyl 2-(trimethylammonio)ethylphosphonate [<sup>31</sup>P NMR

(CDCl<sub>3</sub>) δ = 23.3 ppm]. The solid product was filtered, washed and dried using a rotary evaporator. This product was deprotected by reaction with 5 molar equiv. of TMSBr in dichloromethane at 25 °C overnight to give TMAEP [<sup>31</sup>P NMR (CDCl<sub>3</sub>) δ = 16.5 ppm]. The <sup>1</sup>H-decoupled <sup>31</sup>P-NMR spectra of TMAEP and its 1-hydroxy derivative produced enzymatically in this study exhibit a triplet splitting (see main text Figures 3 and 5 and Figure S7), by contrast to the singlet observed in the corresponding spectra of both DMAEP and 2-AEP. We presume that the splitting arises from the *I* = 1 <sup>14</sup>N nucleus on C2, three bonds removed from the <sup>31</sup>P of the phosphonate. The precise reason that the splitting is present uniquely in the *N*-trimethylammonium-containing compound is not clear, but it may relate to the effect of this substituent on the dynamics of rotation about the C1–C2 bond. Alternatively, a persistent through-space interaction of the trimethylammonium and phosphonate ions may engender the unexpected coupling. The final products were extracted twice with 25 mL of a solution of 3 molar equiv. of ammonium acetate, partially dried using a rotary evaporator followed by lyophilization, and stored at -20 °C.



**Scheme S1.** Synthesis of TMAEP and DMAEP

**Table S1.** Target ion  $m/z$  ratios detected in positive mode.

<b>compound</b>	<b><math>m/z</math></b>	<b><math>m/z +16</math></b>
TMAEP	168	184
DMAEP	154	170
PC	184	200
$\gamma$ bb	146	162
$\gamma$ bb-3	132	148
choline	104	120
glycine betaine	118	
glycine betaine aldehyde	102	

**Table S2.** Target ion  $m/z$  ratios detected in positive mode.

<b>compound</b>	<b><math>m/z</math></b>	<b><math>m/z +16</math></b>
2-AEP	126	142
taurine	168	184
GABA	104	120
$\beta$ -Alanine	90	106
dimethyl glycine	104	
glycine	76	

**Table S3.** Data collection and refinement statistics for the X-ray crystal structures<sup>1</sup> of TmpA in the absence of substrate/product and in the substrate-bound and product-bound complexes.

	<b>TmpA•Fe(II)• 2OG</b>	<b>TmpA•Fe(II)• 2OG•TMP</b>	<b>TmpA•Fe(II)• 2OG•TMO</b>
<b>Data Collection</b>			
Wavelength	0.97857 Å	0.97857 Å	0.97857 Å
Space Group	<i>P4<sub>1</sub>2<sub>1</sub>2</i>	<i>P4<sub>1</sub>2<sub>1</sub>2</i>	<i>P4<sub>1</sub>2<sub>1</sub>2</i>
Cell Dimensions			
a, b, c (Å)	87.051, 87.051, 220.696	86.901, 86.901, 220.613	87.564, 87.564, 221.119
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	50.00-1.73 (1.76- 1.73)	50.00-1.70 (1.73- 1.70)	50.00-1.78 (1.81- 1.78)
R <sub>merge</sub>	0.065 (0.871)	0.052 (0.633)	0.083 (0.467)
< I/σ >	26.4 (2.2)	34.9 (3.1)	25.4 (3.9)
CC <sub>1/2</sub>	0.776	0.851	0.914
Completeness (%)	98.8 (99.9)	99.6 (100.0)	99.8 (99.7)
Redundancy	8.2 (7.9)	8.1 (7.2)	8.3 (7.7)
<b>Refinement</b>			
Resolution (Å)	1.73	1.70	1.78
No. Reflections	83209	88905	79060
R <sub>work</sub> / R <sub>free</sub>	0.194/0.214	0.200/0.220	0.197/0.217
No. atoms			
Protein	5880	5852	5890
Ion/Ligand	32	27	13
Water	464	521	483
B-Factors			
Protein	21.4	15.6	21.5
Ion/Ligand	34.4	11.2	13.9
Water	24.9	21.1	26.3
r.m.s. deviations			
Bond lengths (Å)	0.006	0.006	0.006
Bond angles (°)	1.17	1.20	1.20
Ramachandran			
Outliers (%)	2 (0.27)	0 (0.00)	1 (0.13)
Molprobability Score (%-tile)	0.71 (100)	0.79 (100)	0.65 (100)

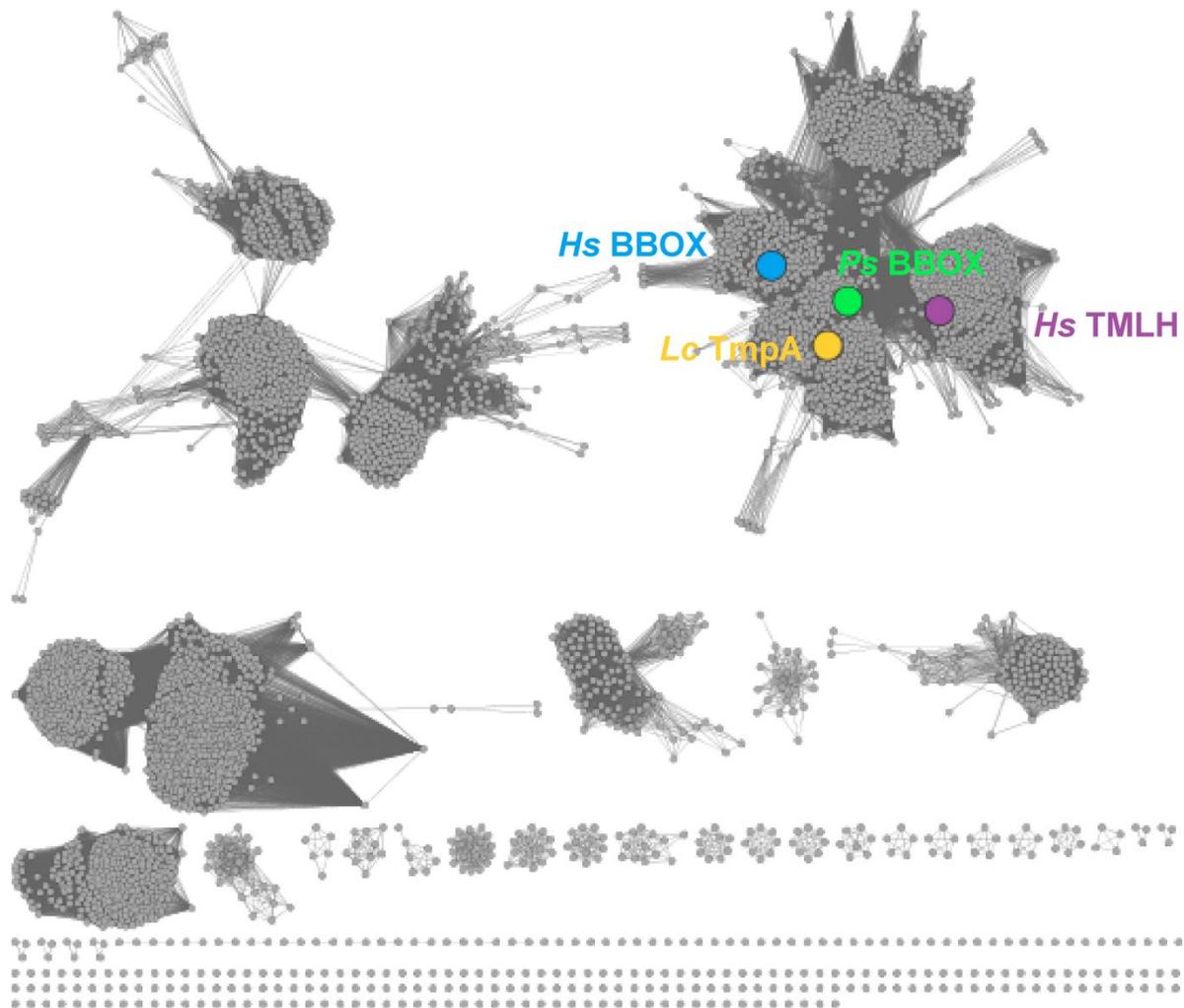
<sup>1</sup>A single crystal was used to solve each structure.

\*Highest resolution shell is shown in parenthesis

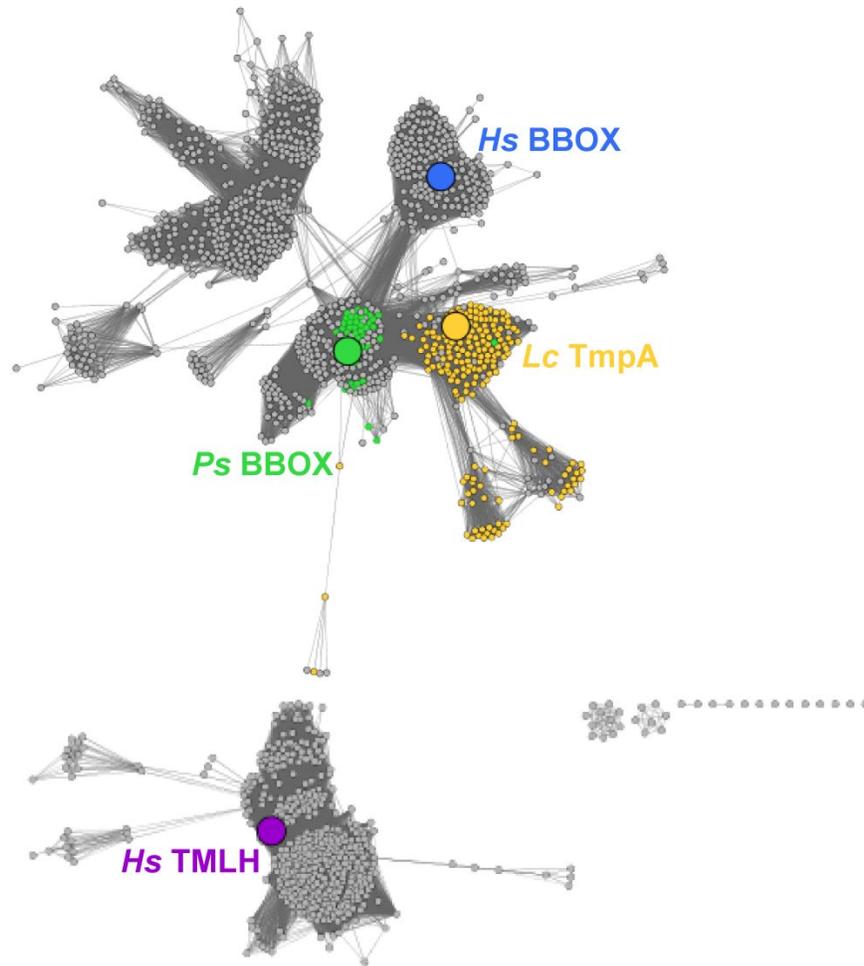
**Table S4.** Data collection and refinement statistics for the X-ray crystal structure of TmpB.

	TmpB•Fe(II)•TMO	TmpB•Fe(II)•TMO (7.2 keV)
<b>Data Collection</b>		
Wavelength	0.97857 Å	1.72200 Å
Space Group	C222 <sub>1</sub>	C222 <sub>1</sub>
Cell Dimensions		
a, b, c (Å)	70.110, 151.322, 135.833	80.760, 66.653, 62.889
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	50.00-1.73 (1.76- 1.73)	50.00-2.54 (2.58- 2.54)
R <sub>merge</sub>	0.067 (0.957)	0.056 (0.190)
< I/ $\sigma$ >	25.9 (2.1)	13.8 (6.7)
CC <sub>1/2</sub>	0.874	0.967
Completeness (%)	99.9 (100.00)	78.3 (80.6)
Redundancy	7.4 (7.3)	3.3 (3.2)
<b>Refinement</b>		
Resolution (Å)	1.72	
No. Reflections	71844	
R <sub>work</sub> / R <sub>free</sub>	0.224/0.251	
No. atoms		
Protein	5882	
Ion/Ligand	32	
Water	261	
B-Factors		
Protein	23.9	
Ion/Ligand	20.8	
Water	26.8	
r.m.s. deviations		
Bond lengths (Å)	0.006	
Bond angles (°)	1.07	
Ramachandran		
Outliers (%)	1 (0.14)	
Molprobability Score (%-tile)	0.75 (100)	

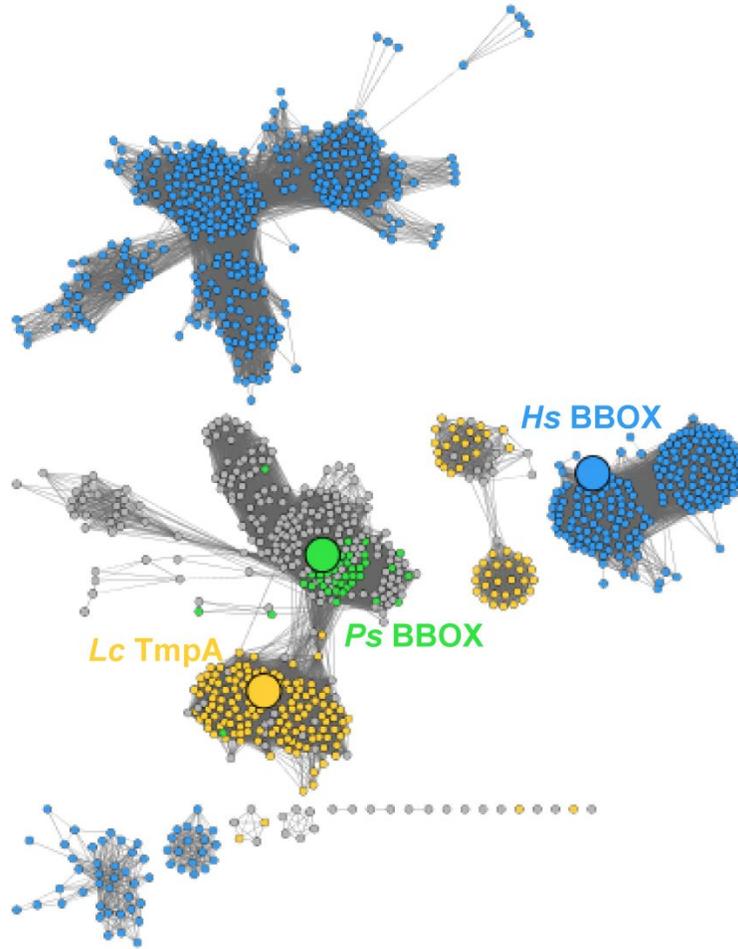
\*Highest resolution shell is shown in parenthesis.



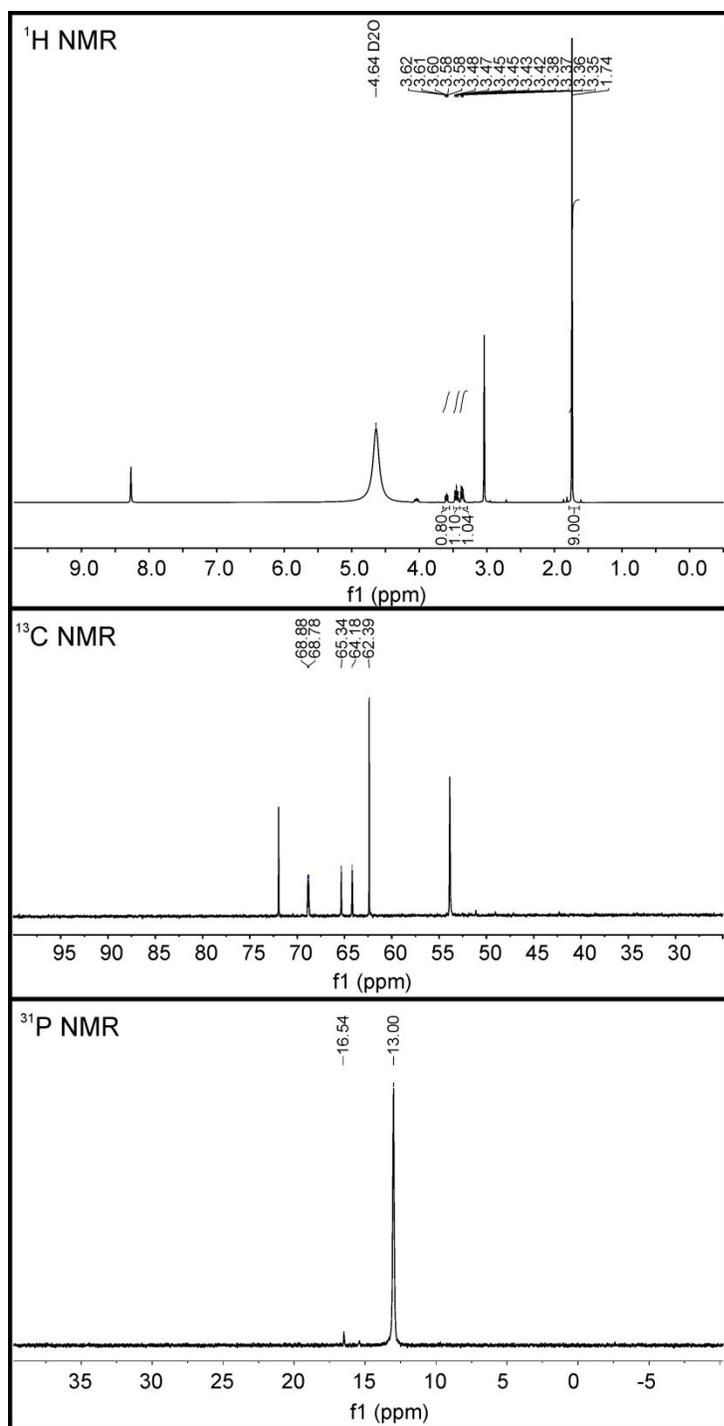
**Figure S1.** SSN constructed from sequences with the IPR003819 domain (length > 350 amino acids) and sequences resulting from BLAST searches of the NCBI and IMG databases using *LcTmpA* as the sequence query, as described in the Experimental section. Nodes represent individual or multiple sequences with > 90% identity. Edges between nodes represent pairwise alignment scores with an *E*-value <  $1 \times 10^{-50}$ . The cluster containing the nodes for *Hs* BBOX (blue), *Hs* TMLH (purple), *Ps* BBOX (green) and *Lc* TmpA (yellow) was isolated from the rest of the network for further analysis (see Figure S2, S3, main text Figure 2).



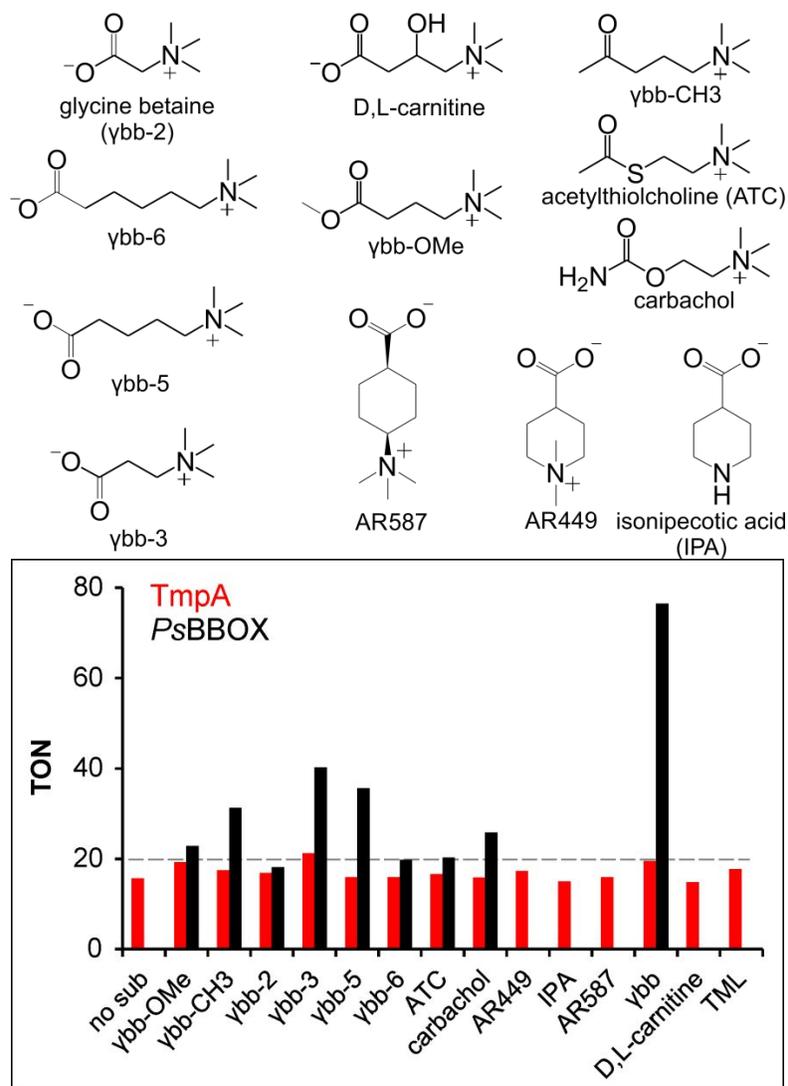
**Figure S2.** SSN of the isolated “BBOX-TMLH” cluster described in Figure S1 applying an  $E$ -value threshold of  $< 1 \times 10^{-65}$  for edges. At this threshold, the node corresponding to *Hs* TMLH (purple) separates into a presumably isofunctional cluster and the BBOX cluster shows sub-clustering based on domain [eukaryotic *Hs* BBOX (blue) versus bacterial *Ps* BBOX (green)] and genomic context [*Ps* BBOX (green) versus *Lc* TmpA (yellow)]. Green nodes represent sequences encoded within a  $\gamma$ bb-degradation operon (see main text Figure 2). Yellow nodes represent sequences encoded in an operon with an HD-domain protein (see main text Figure 2).



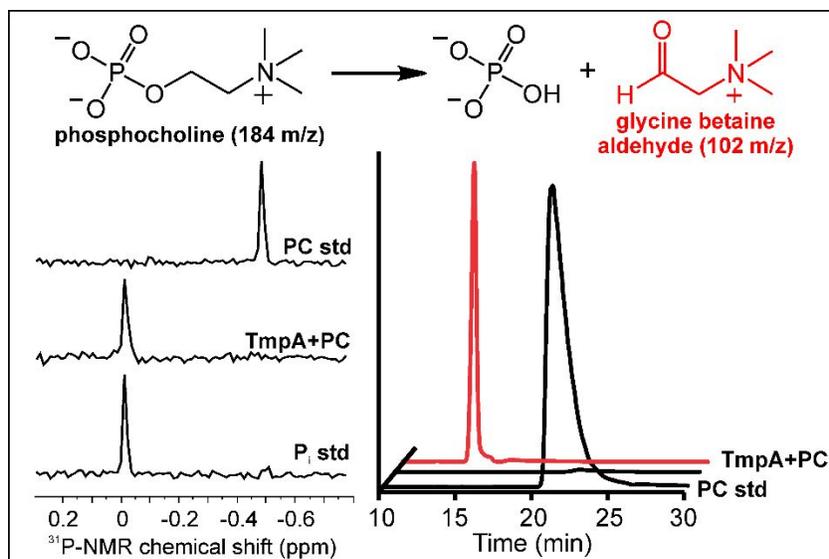
**Figure S3.** SSN of the BBOX cluster (derived from Figure S2) applying an  $E$ -value threshold of  $< 1 \times 10^{-71}$  for edges. At this threshold, the eukaryotic clusters (blue) have become isolated from the bacterial clusters. The presumptively isofunctional clusters containing the TmpA-like nodes (yellow) and BBOX-like nodes (green) have become more isolated, with only three nodes connecting the two major clusters. The bacterial clusters are shown in main text Figure 2.



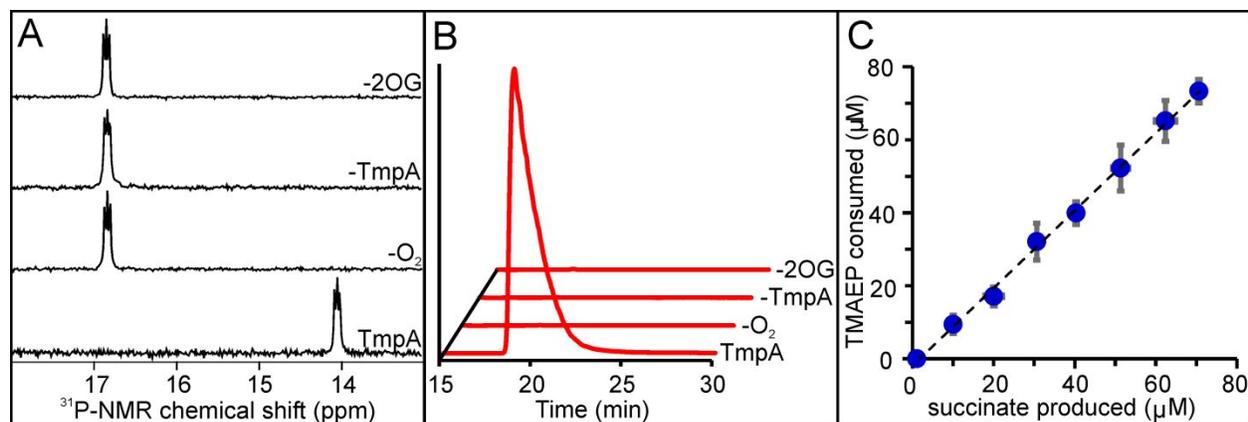
**Figure S4.** NMR ( $^1\text{H}$ , top;  $^{13}\text{C}$ , middle;  $^{31}\text{P}$ , bottom) characterization of the purified (*R*)-OH-TMAEP prepared from TMAEP by an enzymatic reaction with TmpA, as described in the Experimental section. The  $^1\text{H}$ -decoupled  $^{31}\text{P}$  spectrum of (*R*)-OH-TMAEP reproducibly shows a triplet signal (see also main text Figure 3 and Figure S7), similar to that observed for TMAEP, instead of the expected singlet.



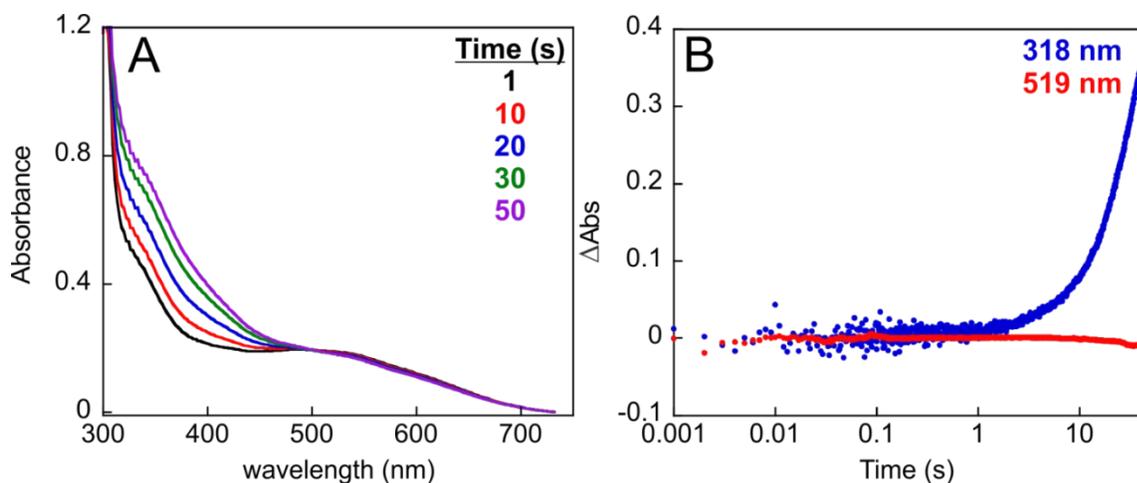
**Figure S5.** *Lc* TmpA and *Ps* BBOX activities toward  $\gamma$ bb analogues. Succinate production was monitored by LC-MS in a reaction containing 0.01 mM TmpA or *Ps* BBOX, 0.02 mM  $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$ , 0.2 mM L-ascorbate, 1 mM 2OG and 1 mM analogue. Uncoupled succinate production in the absence of any substrate was found to be  $17 \pm 2$  turnovers (TON) after a 4-h reaction. Whereas *Ps* BBOX is promiscuous with substrate analogues<sup>7</sup>, none of the compounds in this panel stimulated succinate production by TmpA. Additionally, LC-MS analysis did not reveal any new peaks with  $m/z +16$  relative to those of the substrates that would correspond to hydroxylated products.



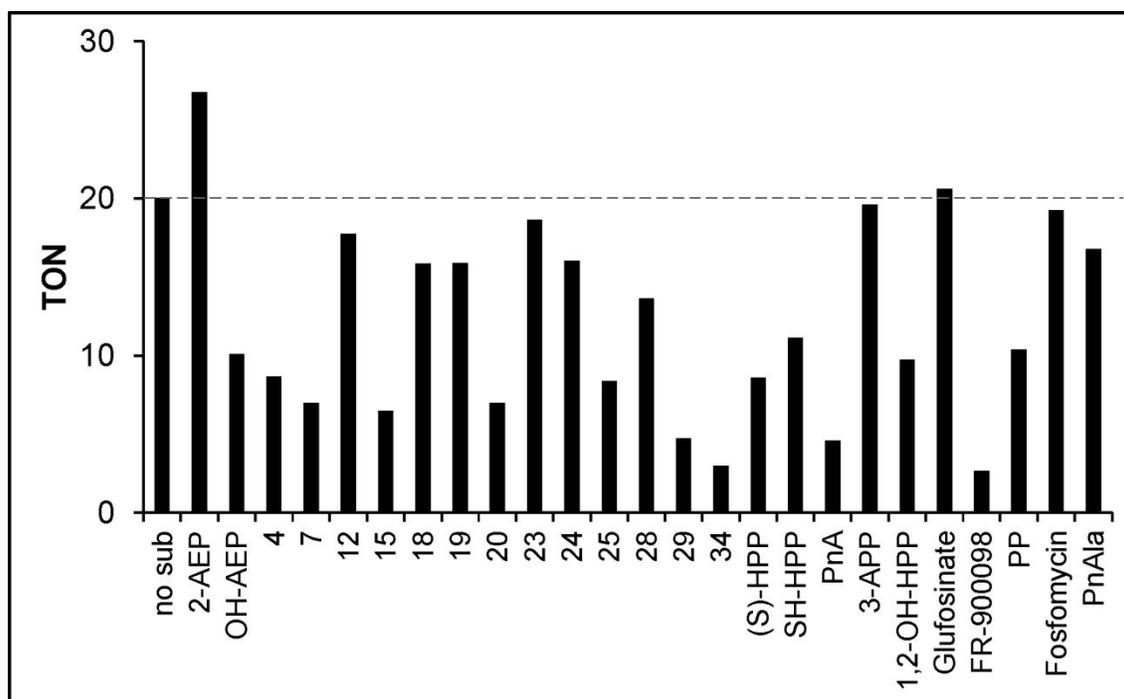
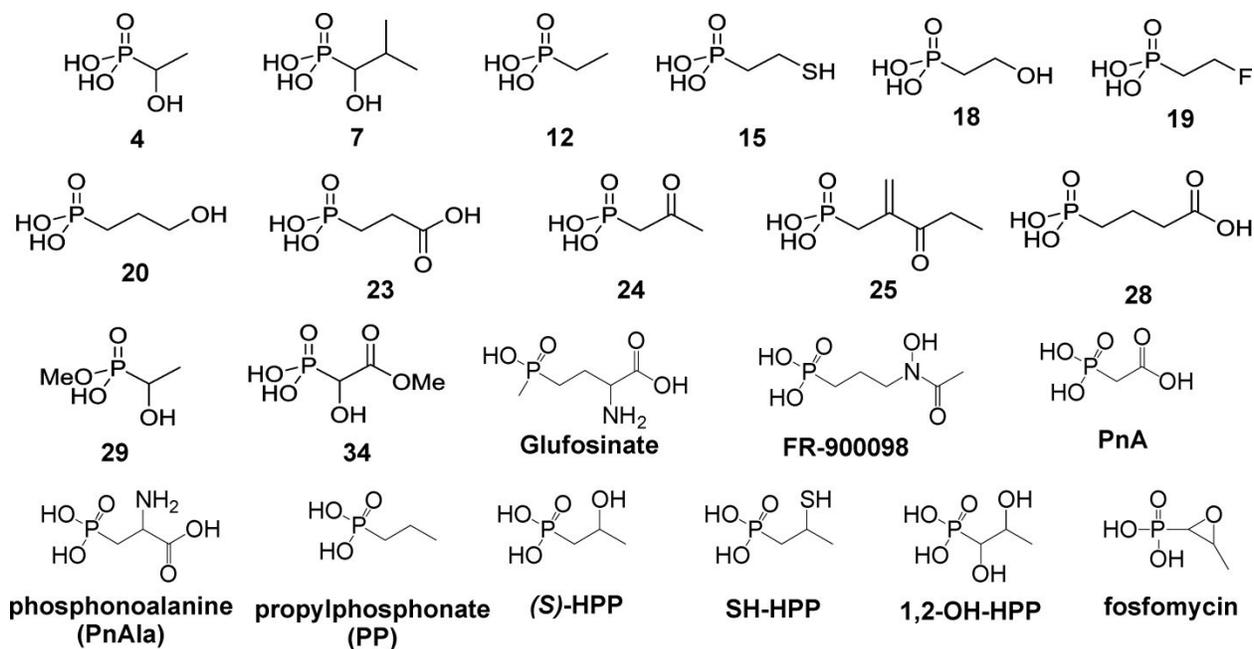
**Figure S6.** *Lc* TmpA activity toward phosphocholine (PC). (Left)  $^{31}\text{P}$ -NMR spectrum of a reaction containing 0.01 mM TmpA, 0.02 mM  $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$ , 0.2 mM L-ascorbate, 3 mM 2OG and 2 mM PC after a 4-h incubation with air flushing. (Right) LC-MS chromatograms of the same sample as in panel A, monitoring PC (black) and glycine betaine aldehyde (red).



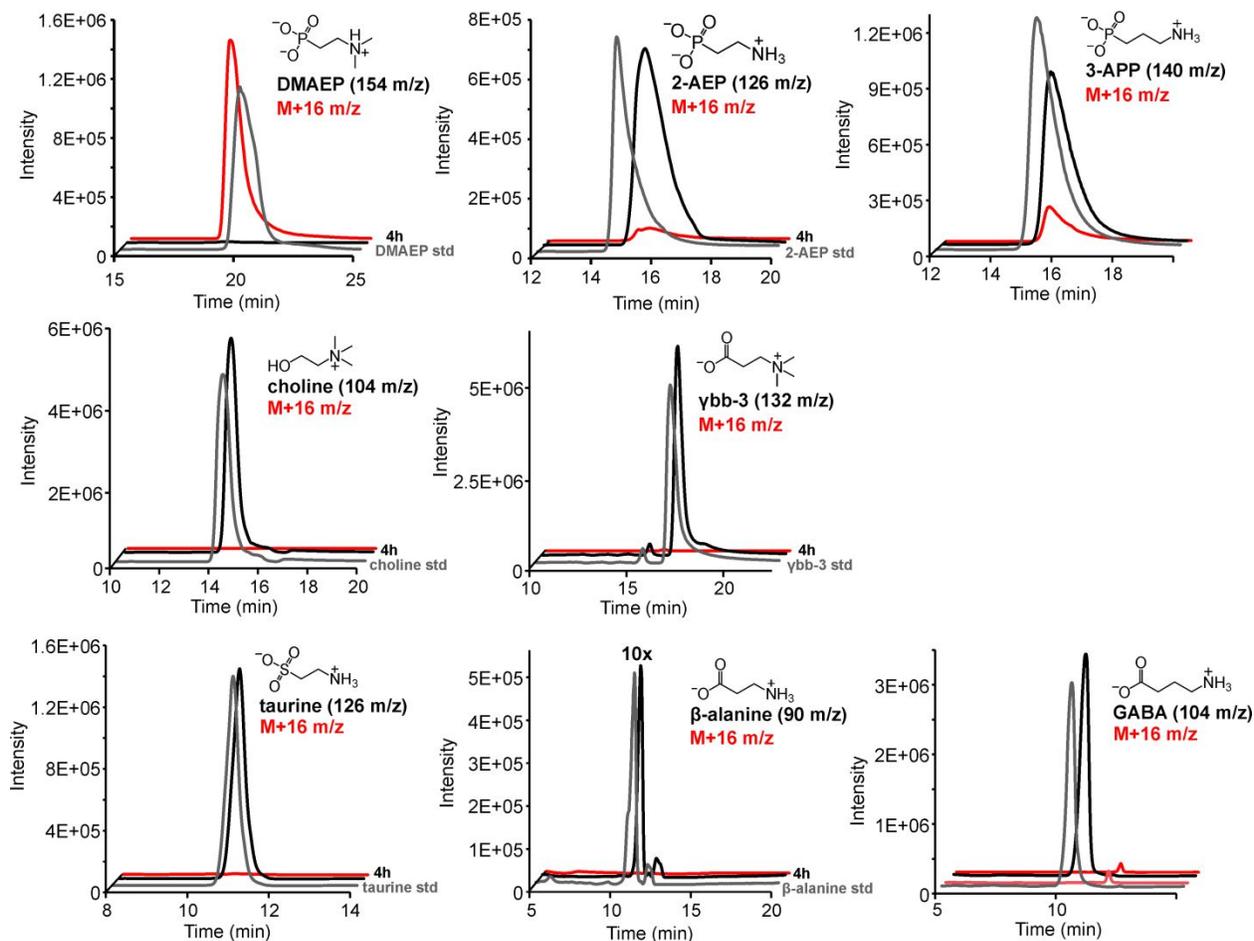
**Figure S7.** Requirements for TMAEP hydroxylation by TmpA. (A)  $^{31}\text{P}$ -NMR spectra of a reaction containing 0.01 mM TmpA, 0.02 mM  $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$ , 0.2 mM L-ascorbate, 3 mM 2OG and 2 mM TMAEP after a 4-h incubation with air flushing, and parallel reactions omitting 2OG, TmpA, or  $\text{O}_2$ . (B) LC-MS single-ion-monitoring chromatograms monitoring OH-TMAEP product formation ( $m/z = 184$ ) in samples parallel to those in panel A. (C) Determination of the "coupling ratio," defined as the ratio of TMAEP consumed to succinate produced in reactions with limiting 2OG; the slope of the linear fit is  $1.0 \pm 0.1$ .



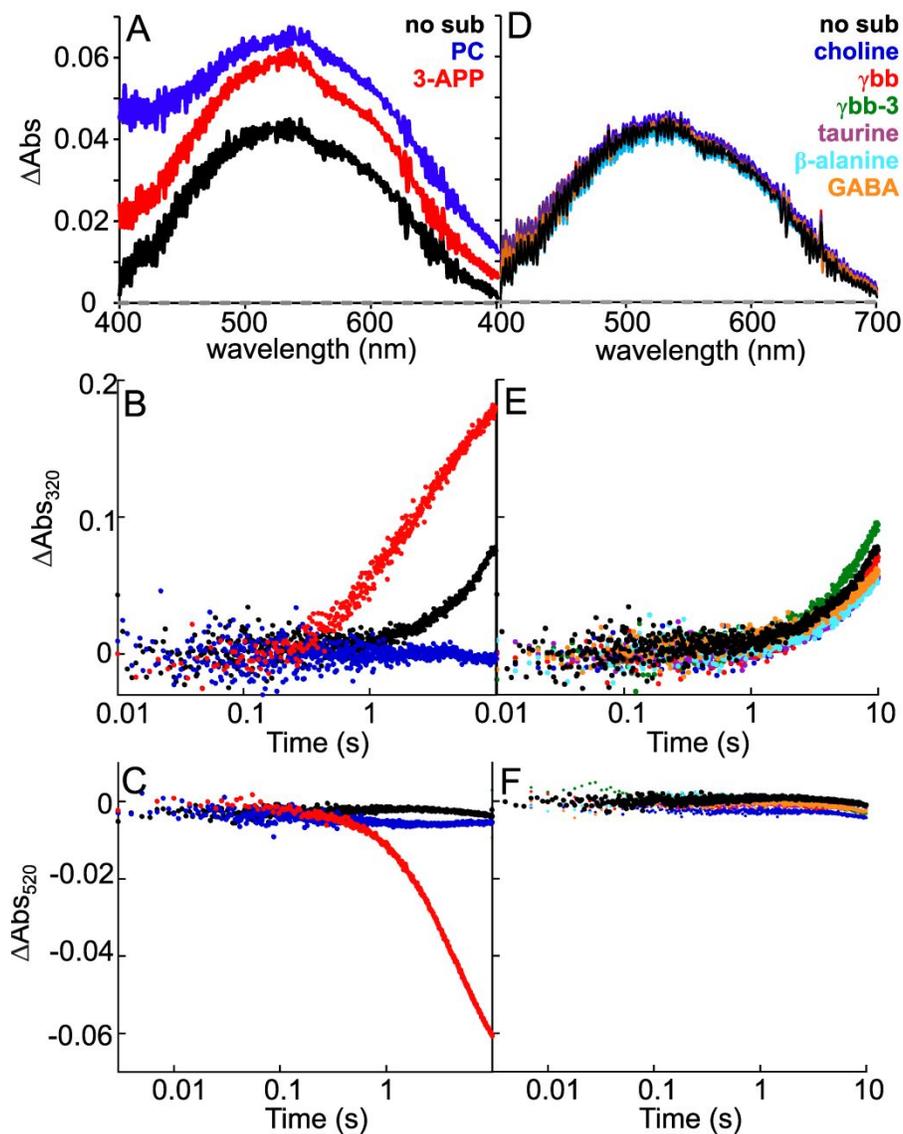
**Figure S8.** Reaction of the TmpA•Fe(II)•2OG complex with O<sub>2</sub> in the absence of a primary substrate. (A) Stopped flow absorption spectra at select time points upon mixing at 5 °C of a solution of 2 mM TmpA, 1.9 mM (NH<sub>4</sub>)<sub>2</sub>Fe(SO<sub>4</sub>)<sub>2</sub> and 10 mM 2OG with an equal volume of O<sub>2</sub>-saturated 50 mM HEPES buffer, pH 7.5 (~1.8 mM O<sub>2</sub>). (B) Kinetic traces at 318 nm (blue) and 519 nm (red) obtained from the time-dependent spectra shown in panel A.



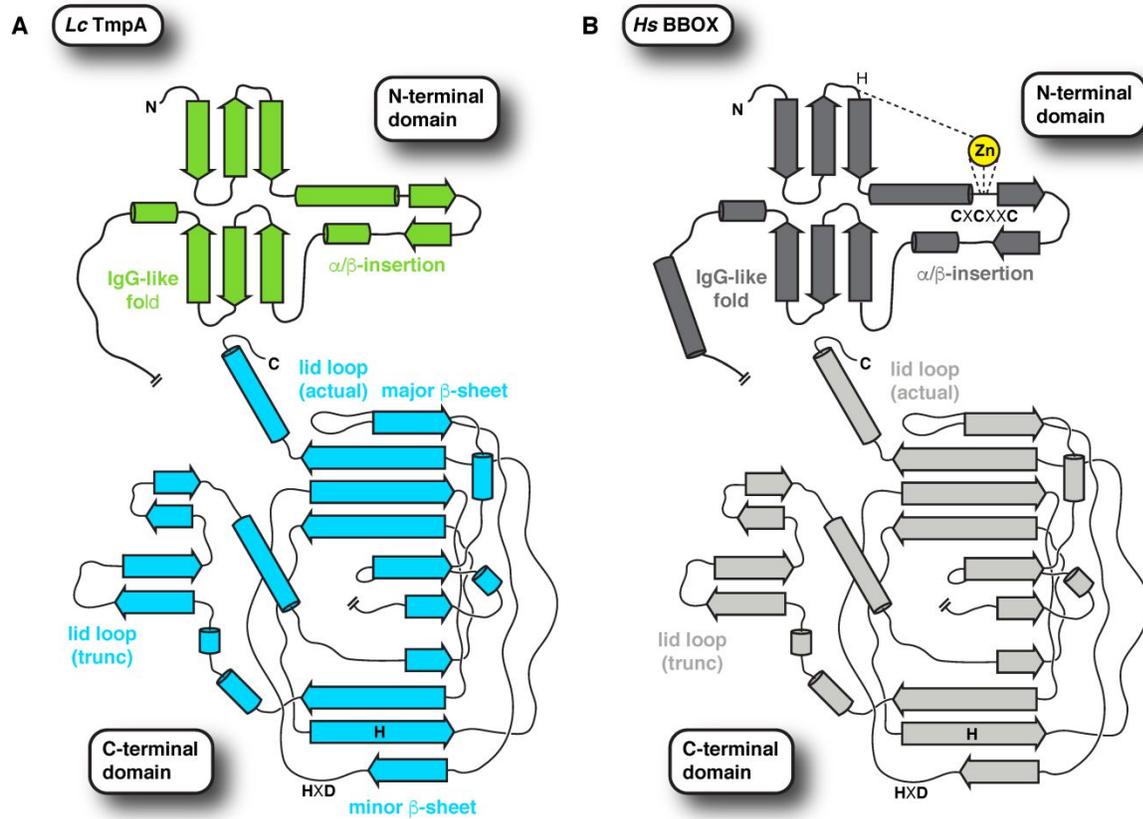
**Figure S9.** *Lc* TmpA activity toward phosphonate analogues. Succinate production was monitored by LC-MS from a reaction of 0.01 mM *Lc* TmpA, 0.02 mM  $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$ , 0.2 mM L-ascorbate, 1 mM 2OG and 1 mM analogue. Uncoupled succinate production in the absence of any substrate was found to be  $17 \pm 2$  turnovers (TON) after a 4-h reaction. Succinate production by *Lc* TmpA was not detectably stimulated in the presence of any analogues lacking an amino functionality. Additionally, LC-MS analysis did not show any new peaks with  $m/z +16$  relative to those of the substrates that would correspond to hydroxylation products.



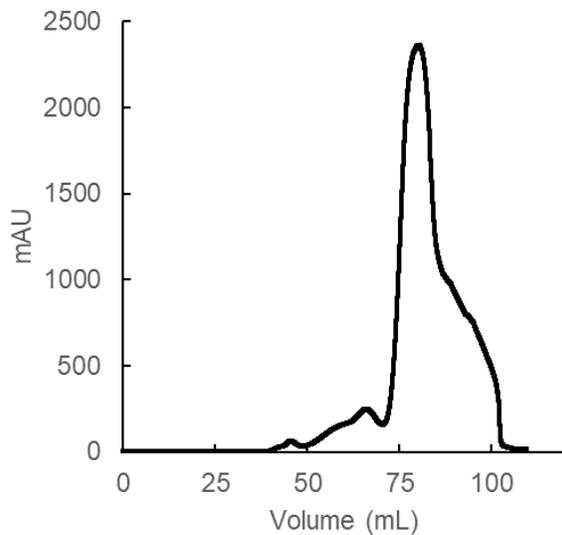
**Figure S10.** LC-MS chromatograms of samples after a 4-h reaction of 0.01 mM TmpA, 0.02 mM  $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$ , 0.2 mM ascorbate, 1 mM 2OG, and 1 mM substrate analogue. Substrate standards are shown in grey. TmpA was able to hydroxylate DMAEP, 2-AEP (as described in the main text) and the unnatural product, 3-aminopropylphosphonate (3-APP), achieving 15 turnovers of the latter after the 4-h reaction. No substrate consumption nor hydroxylated product ( $m/z = +16$  relative to the reactant) was observed for the remaining analogues.



**Figure S11.** Absorption spectroscopic analysis of substrate binding to TmpA•Fe(II)•2OG and promotion of O<sub>2</sub> activation. (A and D) Absorption spectra of TmpA•Fe(II)•2OG in the absence of substrate (black) and in the presence of various substrate analogues. Kinetic traces at 318 nm (B and C) or 519 nm (E and F) upon mixing at 5 °C of a solution of 2 mM TmpA, 1.9 mM (NH<sub>4</sub>)<sub>2</sub>Fe(SO<sub>4</sub>)<sub>2</sub>, 10 mM 2OG and 10 mM substrate analogue with an equal volume of O<sub>2</sub>-saturated 50 mM HEPES buffer, pH 7.5 (~1.8 mM O<sub>2</sub>).

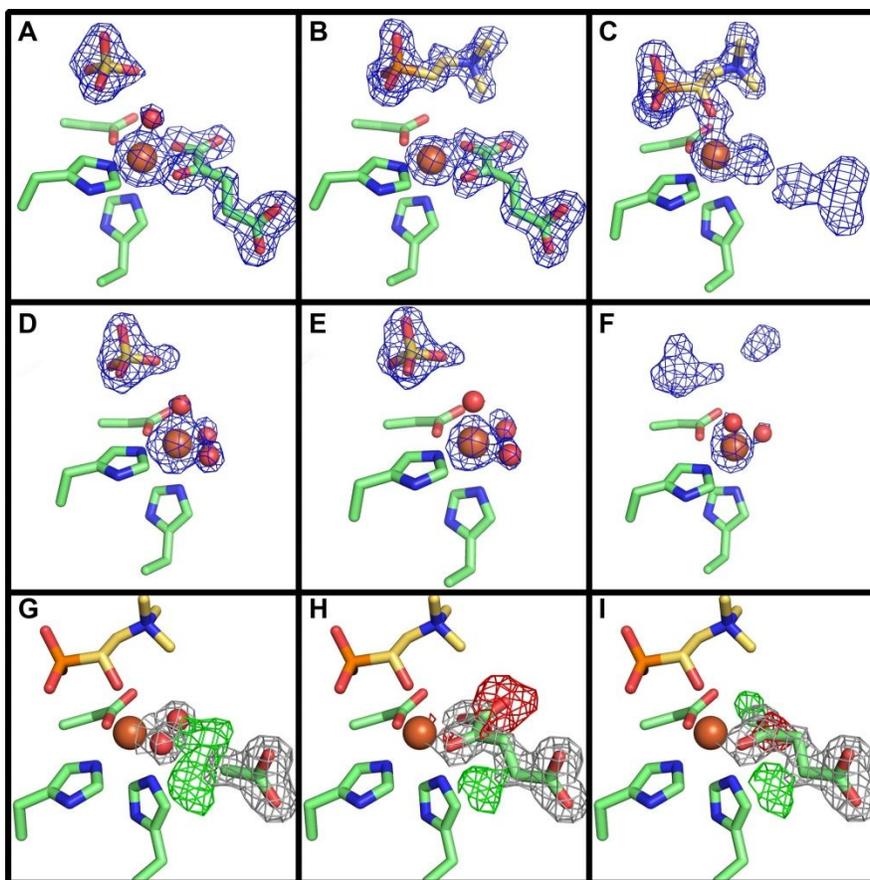


**Figure S12.** Structural topology of (A) *Lc TmpA* and (B) *Hs BBOX* monomers, each consisting of an N-terminal domain and a C-terminal catalytic domain.

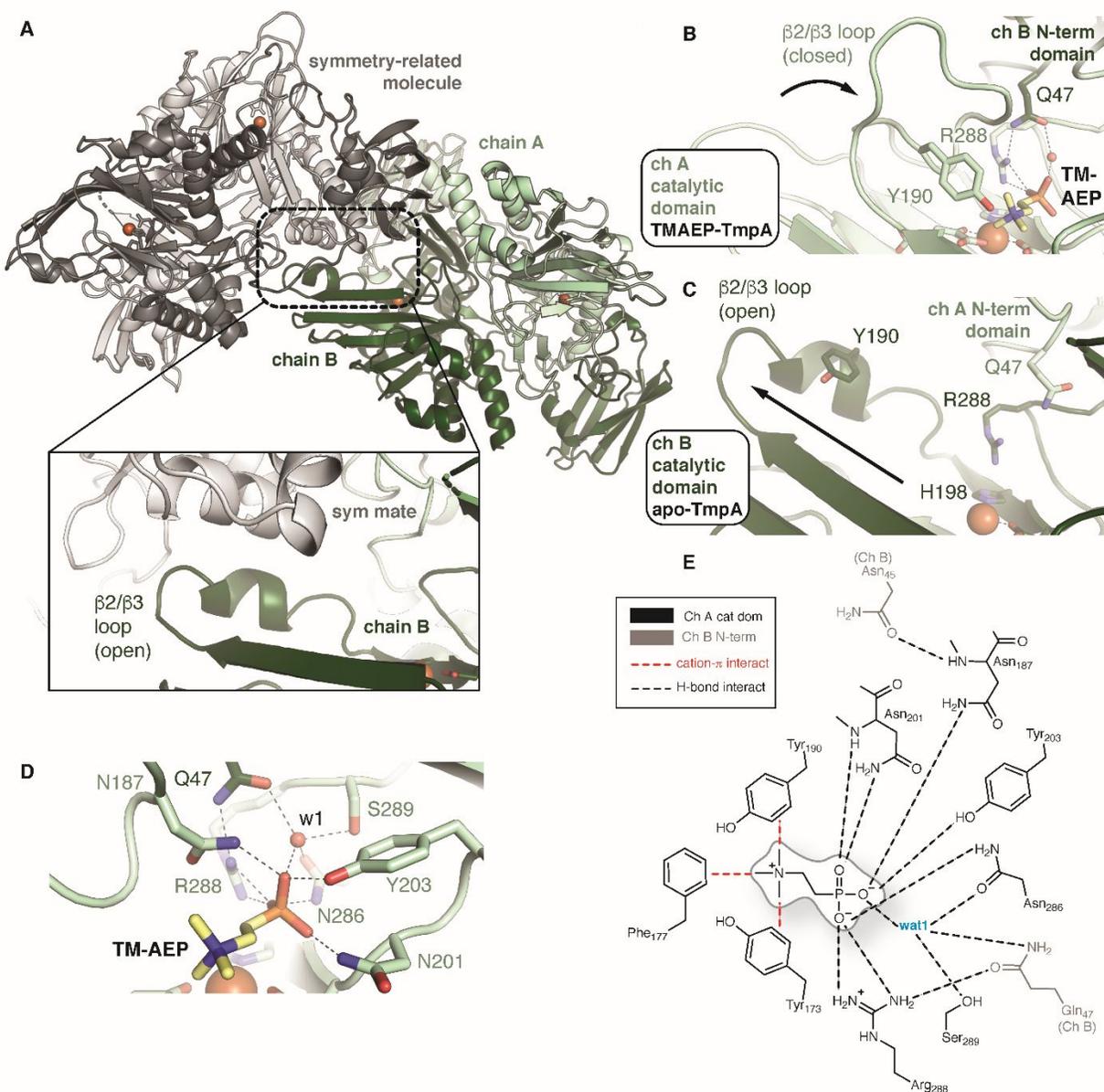


**Figure S13.** Analytical size-exclusion chromatogram of TmpA eluting at  $\sim 78.7$  mL. Calibration of the column ( $V_o = 44.45$  mL,  $V_c = 120.64$  mL) using protein standards (slope =  $-0.365$  Da/mL, y-intercept =  $2.22$  Da) gives an estimated molecular weight of  $\sim 70$  kDa for TmpA. Comparison to the theoretical molecular weight of TmpA ( $41$  kDa<sup>1</sup>) suggests that the TmpA protein is a dimer in solution.

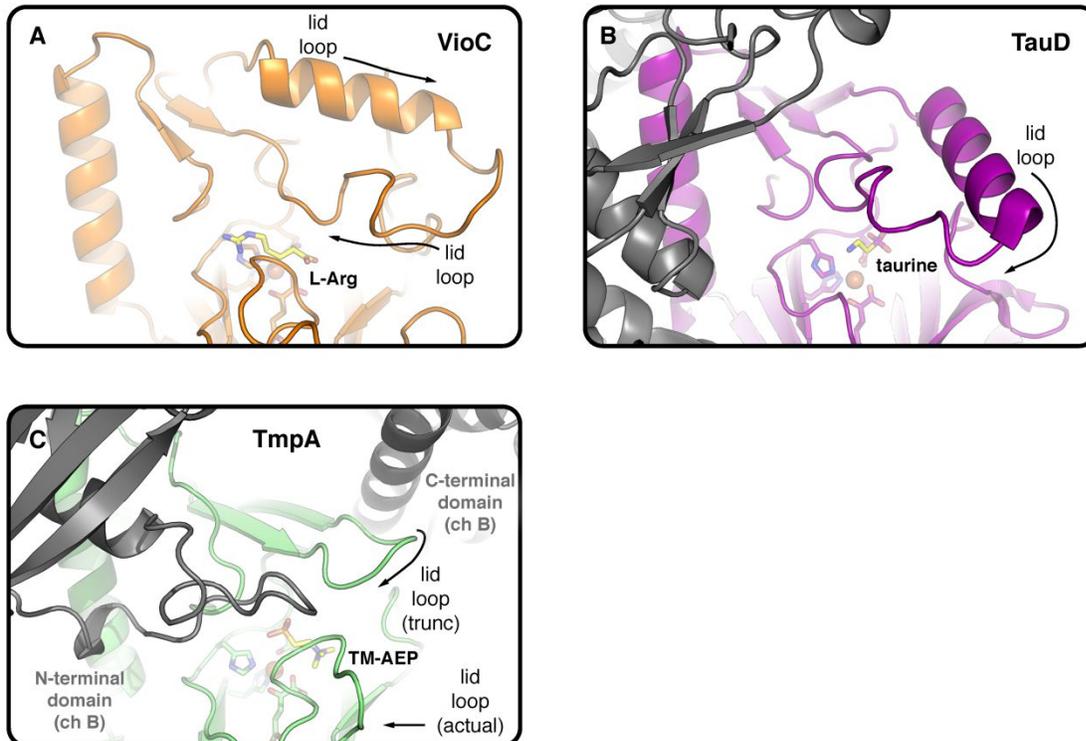




**Figure S15.** Active site models for the TmpA structures. For all panels, the Fe(II) is shown as a sphere, and the coordinating residues His198, Asp200, and His344 are shown as green sticks. The  $F_o - F_c$  difference (green and red mesh) and omit maps (blue mesh) are shown at  $\pm 3.0\sigma$ , and the  $2F_o - F_c$  map (gray mesh) is shown at  $1.5\sigma$ . (A) Chain A of the TmpA ternary complex structure. The  $F_o - F_c$  omit map is shown for Fe(II), 2OG, a water ligand, and a sulfate ion. (B) Chain A of the TmpA quaternary complex structure. The  $F_o - F_c$  omit map is shown for Fe(II), 2OG, and TMAEP (yellow). (C) Chain A of the TmpA product complex structure. The  $F_o - F_c$  omit map is shown for Fe(II), OH-TMAEP (yellow), and the Fe(II) equatorial non-protein ligands. Three models were considered to account for the density for the non-protein ligands: (G) two coordinating water molecules and an acetate molecule resulting from the TMAEP synthesis procedure, (H) 2OG from the crystallization, or (I) succinate resulting from turnover *in crystallo*. None of these options alone accurately model the observed density, suggesting that there is a mixture of states, and, therefore, this density was left unmodeled in the final structure. The negative density for 2OG suggests that the co-substrate likely does not contribute substantially to this mixture, consistent with the observation of substrate hydroxylation. (D) Chain B of the TmpA ternary complex structure. The  $F_o - F_c$  omit map is shown for Fe(II), water ligands, and a sulfate ion. The Fe(II) occupancy is estimated to be  $< 80\%$ . (E) Chain B of the TmpA quaternary complex structure. The  $F_o - F_c$  omit map is shown for Fe(II), water ligands, and a sulfate molecule. The Fe(II) occupancy is estimated to be  $< 65\%$ . (F) Chain B of the TmpA product-bound complex structure. The  $F_o - F_c$  omit map is shown for Fe(II), water ligands, and the partially occupied substrate or product mixture. The Fe(II) occupancy is estimated to be  $< 40\%$ .



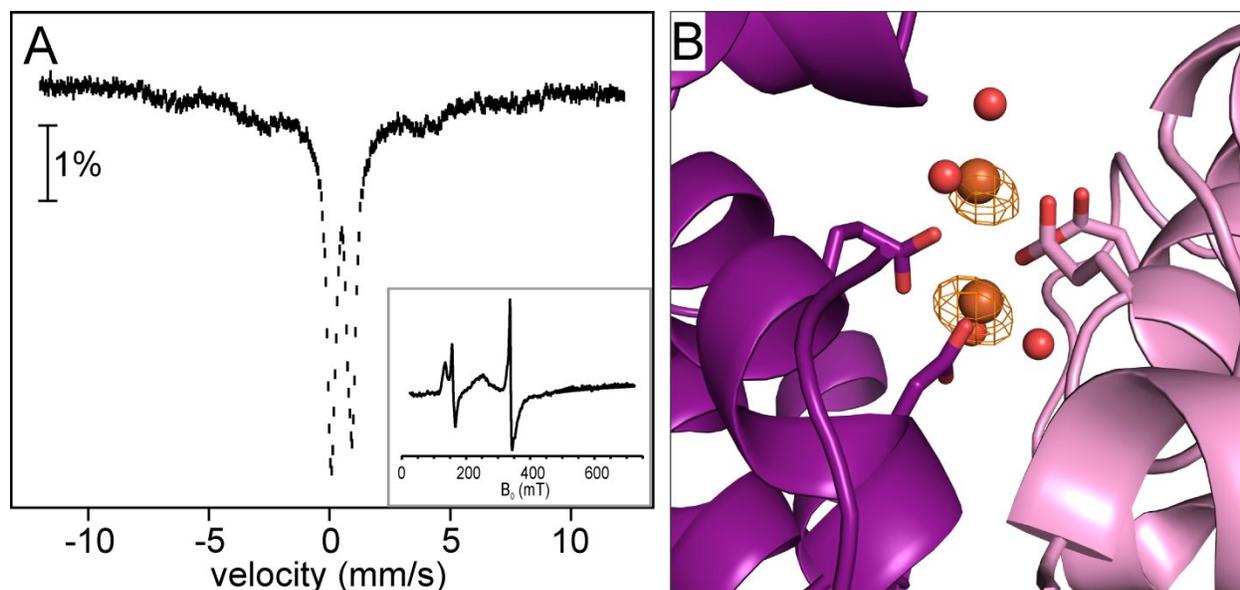
**Figure S16.** Structure of the TmpA•TMAEP complex. (A) Symmetry-mate contacts preventing substrate binding and lid loop closure in chain B, further illustrated by comparison of chain A (B) and chain B (C). (D) Binding pocket for the TMAEP phosphonate group and the hydrogen-bonding network between the chain A C-terminal domain (light green) and the chain B N-terminal domain (dark green) mediated by Q47. (E) Ligand interaction diagram depicting enzyme-substrate contacts.



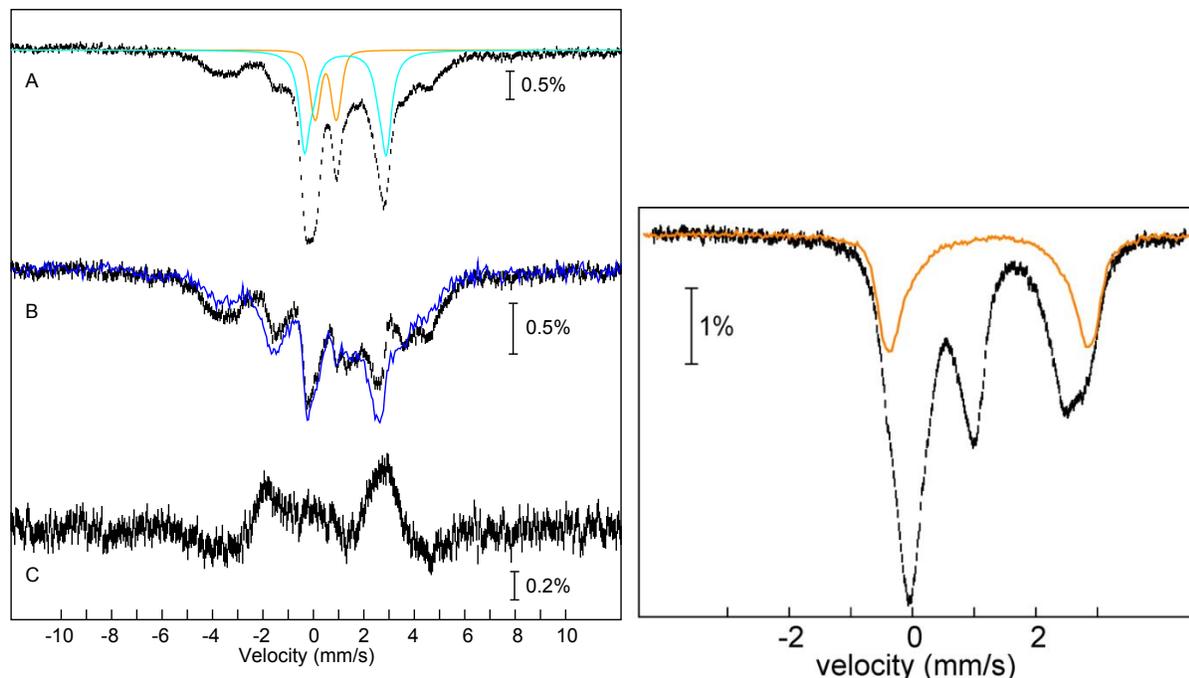
**Figure S17.** Comparison of the active site lid loops in the Fe/2OG enzymes (A) VioC, (B) TauD, and (C) TmpA. The lid-loop region found in other Fe/2OG enzymes is truncated in TmpA (Figure S13). Instead, in TmpA, a different flexible lid loop from the catalytic  $\beta$ -barrel and a loop region of the N-terminal domain from an opposite monomer close over the active site with substrate bound.

TmpB	MSKPDVSKLNRGNIVEFIGGIFDRRGDEEYLGEFVTMAE <b>H</b> MLQGATIAEQNGQPEEIIIVG	60
PhnZ	-----MSLSNSSKVSVLISLLEKSRDLDYIGEAINQ <b>L</b> E <b>H</b> SLQCAYFAQRSGADNEMVLA	54
	. * . . . . * . . : : : : : * : * : * * : . ** * * * : * : : . * : * : : .	
TmpB	ALL <b>H</b> DIGHFTSEFGMFSMDDETEDRY <b>H</b> EEAGAEVLEQFF-PSVITDCVRY <b>H</b> VAAKRYLCAT	119
PhnZ	ALL <b>H</b> DLGHYCNDTSFEDMGYGVW <b>Q</b> <b>H</b> EKVGADYLRGLGFSERVACLIEG <b>H</b> VAAKRYLVSS	114
	*****: ** : : : : . * . . . . ** : . ** : * . : . : : . ***** : :	
TmpB	KPEYFNRLSEASIHSLKLQGGPMDAEEVAE <b>F</b> EKPNPLKQIIAVRYL <b>D</b> EAGKRADMETPDY	179
PhnZ	KASYLKNLSDASRKTLEYQGGPMDEGERRLFEE <b>R</b> EDFKDCLKIRAW <b>D</b> EKGKQTDLKVPGP	174
	* . * : : . * * : * * : : * : ***** * ** : . : * : : : * ** * * : : * : : . *	
TmpB	WHFAPMVQRMVDKHMGA	196
PhnZ	EHYRKMMEEHLENQN-	190
	* : * : : . : : : .	

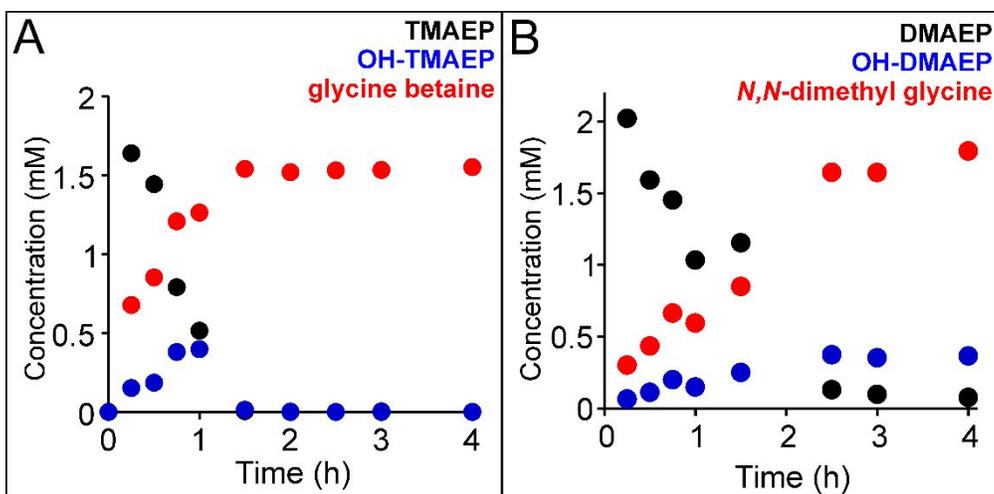
**Figure S18.** Pairwise sequence alignment of PhnZ from *uncultured bacterium HF130\_AEPn\_1* and *LcTmpB*, with the conserved histidine and aspartate residues of the extended HD-domain sequence motif highlighted in red.



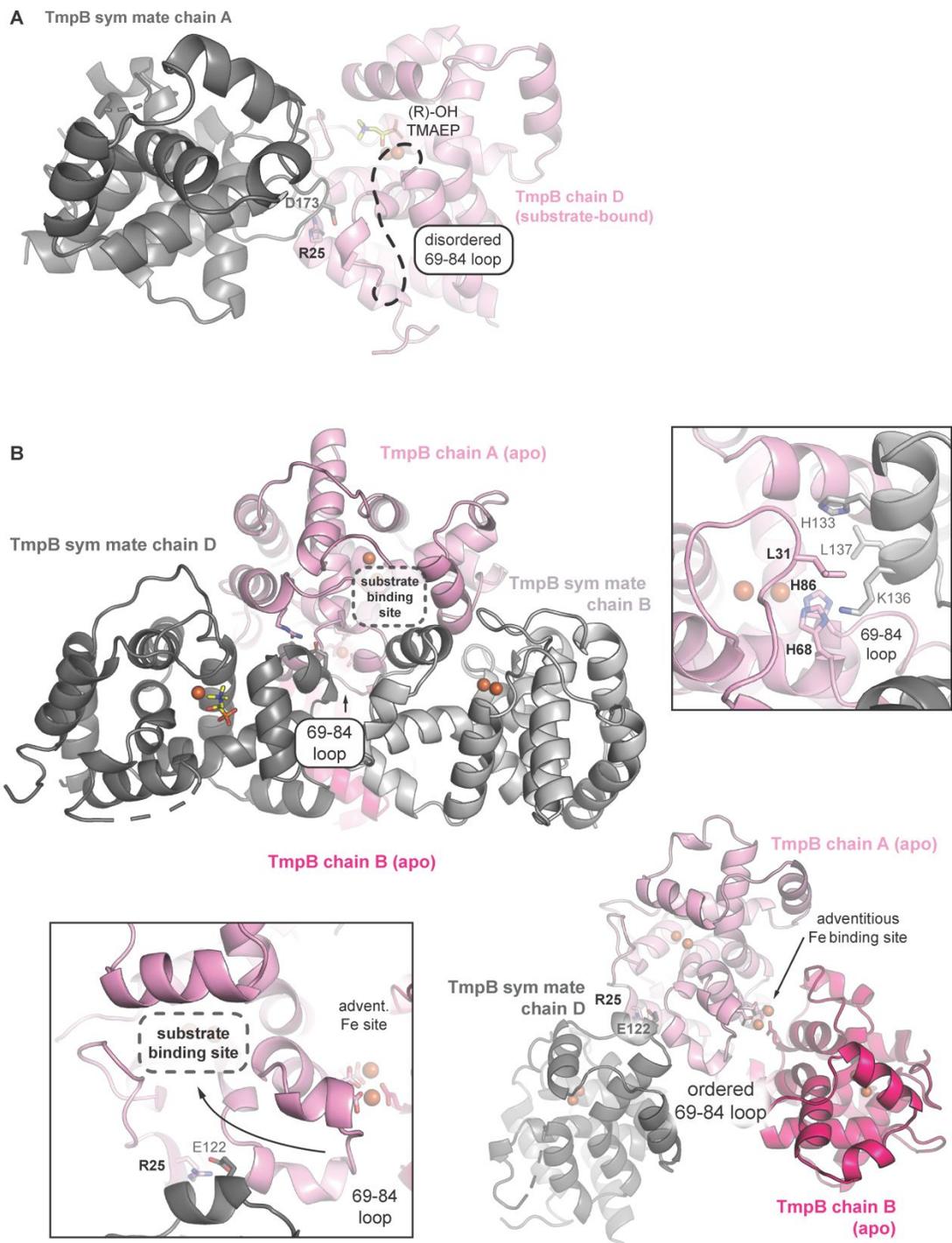
**Figure S19.** Spectroscopic evidence for adventitiously-bound Fe(III) in TmpB. (A) 4.2 K/53 mT Mössbauer spectrum of 2 mM O<sub>2</sub>-free TmpB in the Fe<sub>2</sub>(III/III) state prepared by incubation with 3 mM potassium ferricyanide for 45 min. (Inset) X-band EPR spectrum of a sample of as-isolated, oxidized TmpB (0.25 mM). Experimental conditions: temperature = 10 K, microwave power = 0.2 mW, microwave frequency = 9.479 GHz, modulation amplitude = 1 mT. (C) X-ray crystal structure of TmpB, showing the interface of monomers C and D in the asymmetric unit. Orange mesh depicts an anomalous Fourier electron density map contoured at 5 $\sigma$ . The occupancy of these metals was adjusted to 50% in the final model. Their observation in the crystal structure supports the conclusion that the high-spin Fe(III) signals observed in the Mössbauer and EPR spectra are due to adventitiously-bound iron, not mononuclear Fe(III) in the active site.



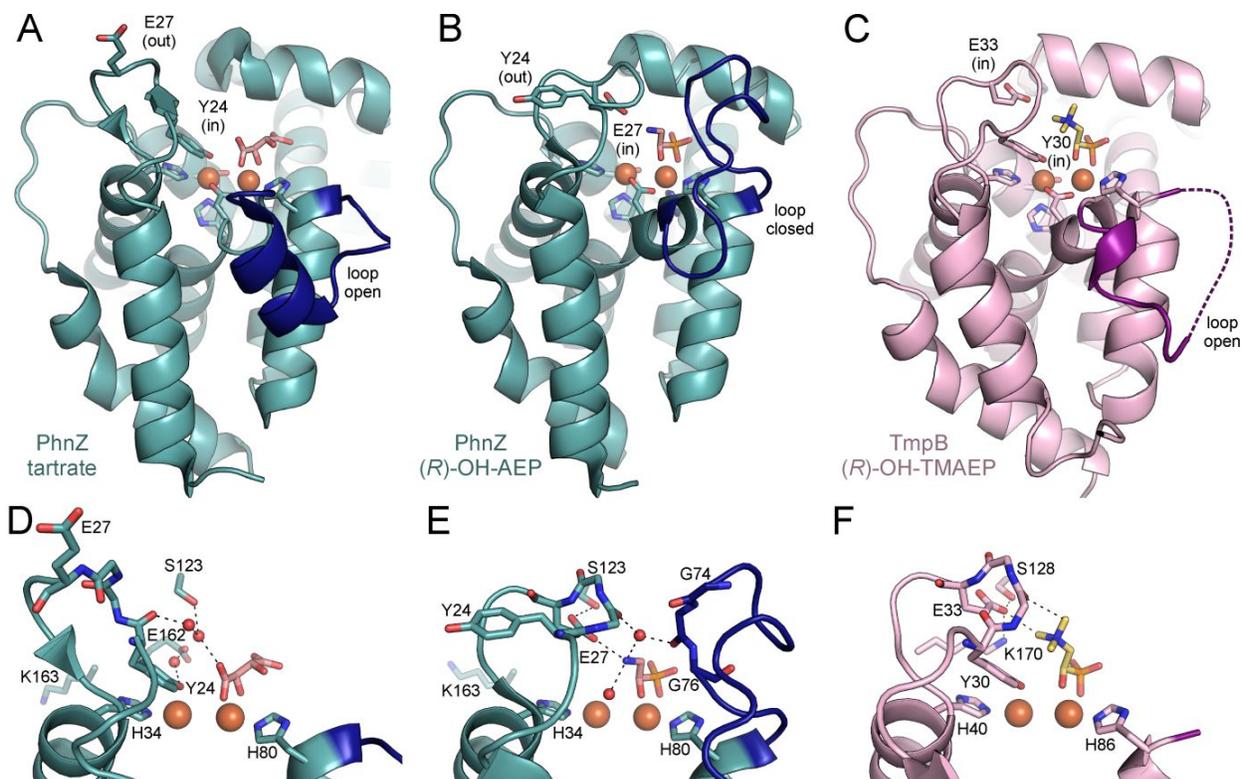
**Figure S20.** Mössbauer spectra of a sample of TmpB enriched in the  $\text{Fe}_2(\text{II/III})$  state. The sample was generated by incubation of a solution of 2 mM  $\text{O}_2$ -free TmpB in the  $\text{Fe}_2(\text{III/III})$  state with 20 mM sodium L-ascorbate for 45 min in the absence of  $\text{O}_2$ . The 4.2K/53mT spectrum with the external magnetic field oriented parallel to the  $\gamma$  beam is shown as black vertical bars in the left panel, A. The spectral contributions from  $\text{Fe}_2(\text{II/II})$  (22% of total intensity) and  $\text{Fe}_2(\text{III/III})$  (13% of total intensity) are shown as turquoise and orange lines, respectively. Removal of these contributions yields the experimental reference spectrum of the TmpB  $\text{Fe}_2(\text{II/III})$  state in a magnetic field of 53 mT applied parallel (left panel, B, black vertical bars). The corresponding reference spectrum of the  $\text{Fe}_2(\text{II/III})$  cluster with the external field oriented perpendicular to the  $\gamma$  beam (left panel, B, blue line) was prepared in an identical fashion. The resulting parallel-minus-perpendicular difference spectrum (left panel, C) of the TmpB  $\text{Fe}_2(\text{II/III})$  state exhibits the field-orientation dependence expected for a cluster with an  $S_{\text{Total}} = 1/2$  electronic ground state. The 120-K/zero-field spectrum of this sample (right panel, black bars) overlaid with 26% of the experimental  $\text{Fe}_2(\text{II/II})$  spectrum collected under identical conditions (orange line), which was then subtracted to give the spectrum shown in the main text Figure 8.



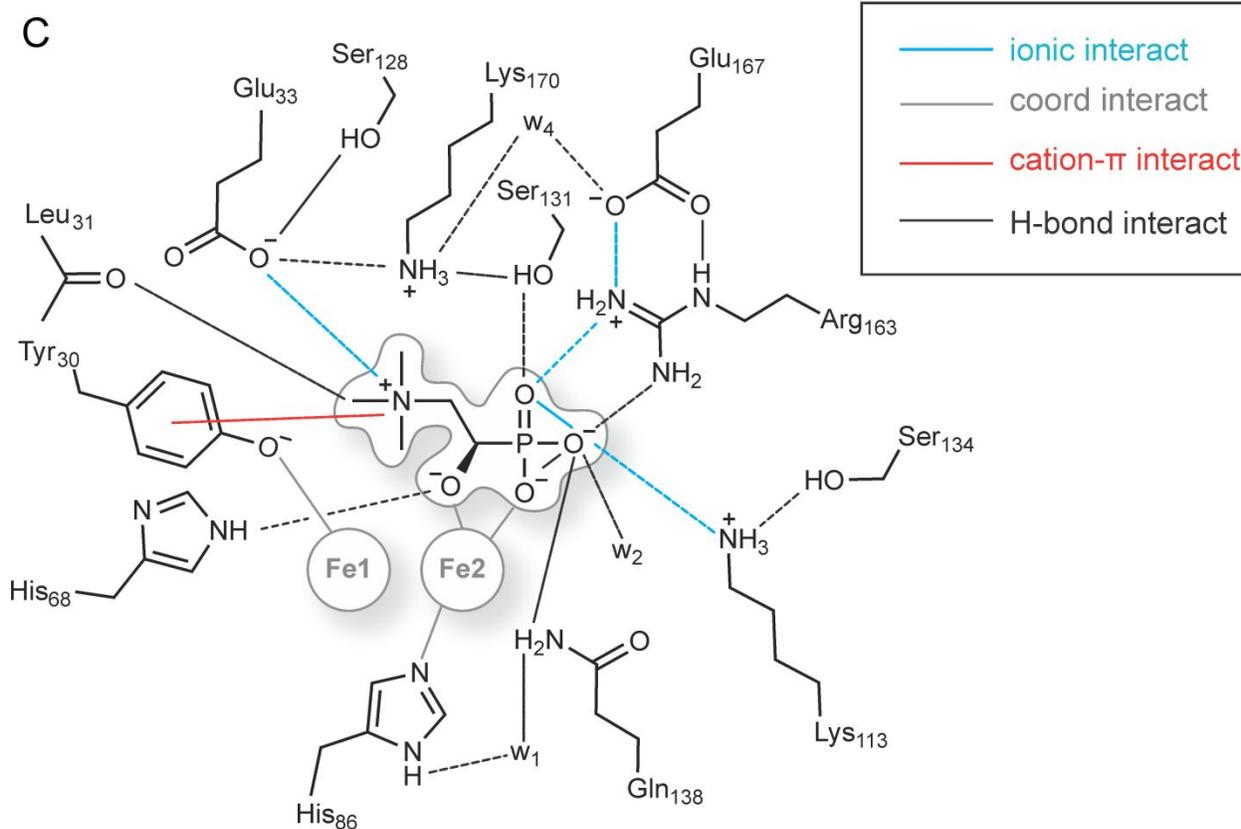
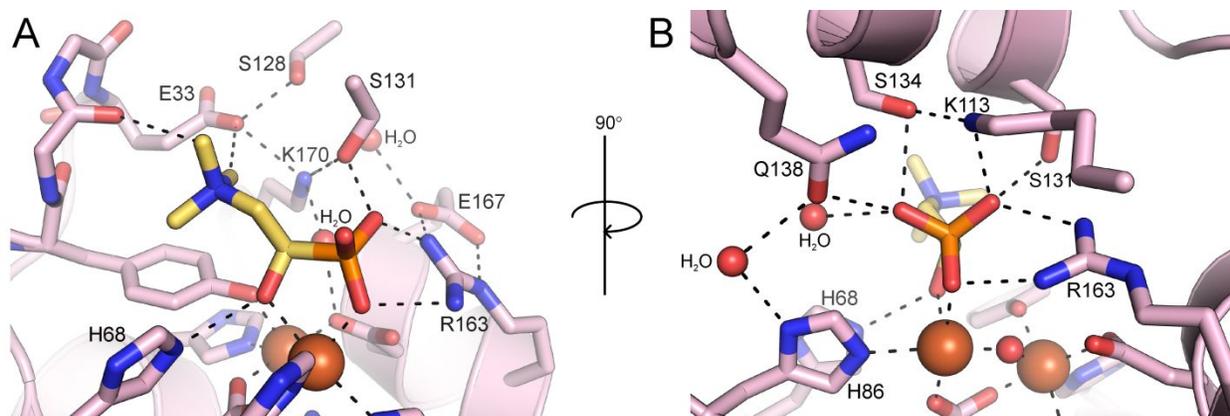
**Figure S21.** Coupled reactions of TmpA and TmpB. Reactions were incubated for 4 h at 3 °C and contained 0.02 mM TmpA, 0.02 mM TmpB, 0.03 mM  $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$ , 0.4 mM L-ascorbate, 6 mM 2OG, and either 2 mM TMAEP (A) or DMAEP (B). Substrates (black), hydroxylated intermediate products (blue) and glycine betaine (A, red) or dimethyl glycine (B, red) were monitored by LC-MS.



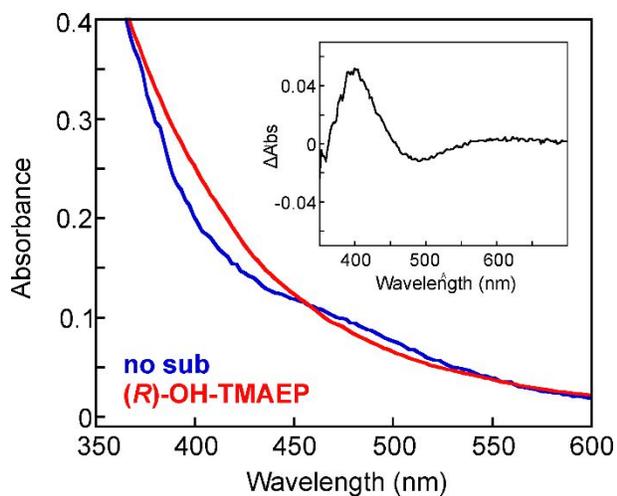
**Figure S22.** (A) A view of specific lattice contacts interactions that induce lid loop disorder and prevent closure in substrate-bound monomers of the TmpB crystal structure. (B) Views of other crystal packing interactions that prevent substrate binding (*top*) or yield the adventitious iron binding site (*bottom right, left*).



**Figure S23.** (A) PhnZ structure with tartrate bound (PDB accession code 4MLM) showing the lid loop (dark blue) consisting of residues 62-78 and the Y<sub>24</sub>XXE<sub>27</sub> loop orientation. (B) PhnZ structure with (*R*)-OH-AEP bound (PDB accession code 4MLN) showing closure of the lid loop (dark blue) and reversal of the Y<sub>24</sub>XXE<sub>27</sub> loop orientation. (C) TmpB structure showing the unmodeled region (dashed line) of the analogous lid loop (dark purple) and the Y<sub>30</sub>XXE<sub>33</sub> loop. Active site zoom of (D) PhnZ•tartrate, (E) PhnZ•(*R*)-OH-AEP, and (F) TmpB•(*R*)-OH-TMAEP depicting interactions between the YxxE loop and the molecule bound in the active site.



**Figure S24.** (A, B) Two views of the (*R*)-OH-TMAEP (yellow) interactions in the TmpB active site. Iron ions are shown as brown spheres. Dashed lines represent hydrogen bonding or electrostatic interactions. (C) Ligand interaction diagram depicting enzyme-substrate interactions.



**Figure S25.** UV/visible absorption spectra of  $\text{Fe}_2(\text{II/III})$  TmpB in the absence of substrate (blue) and the presence of (*R*)-OH-TMAEP (red). The difference spectrum is shown in the inset. The sample contained  $\text{O}_2$ -free 0.125 mM TmpB pre-reduced with 10 mM sodium L-ascorbate and chromatographed on a PD-10 desalting column (GE Healthcare) to remove the sodium L-ascorbate. The protein sample was then incubated with 12.5 mM (*R*)-OH-TMAEP for 1.5 min at 5 °C.

## Supporting Information References

- [1] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, In *The Proteomics Protocols Handbook* (Walker, J. M., Ed.), pp 571-607, Humana Press.
- [2] Fish, W. W. (1988) In *Methods Enzymol.* (James F. Riordan, B. L. V., Ed.), p 357, Academic Press.
- [3] Floch, V., Loisel, S., Guenin, E., Herve, A. C., Clement, J. C., Yaouanc, J. J., des Abbayes, H., Ferec, C. (2000) Cation substitution in cationic phosphonolipids: A new concept to improve transfection activity and decrease cellular toxicity. *J. Med. Chem.* *43*, 4617-4628.
- [4] Li, J., Ni, X. F., Ling, J., Shen, Z. Q. (2013) Syntheses and properties of poly(diethyl vinylphosphonate) initiated by lanthanide tris(borohydride) complexes: Polymerization controllability and mechanism. *J. Polym. Sci. Pol. Chem.* *51*, 2409-2415.
- [5] Lukac, M., Garajova, M., Mrva, M., Devinsky, F., Ondriska, F., Kubincova, J. (2014) Novel fluorinated dialkylphosphonatocholines: Synthesis, physicochemical properties and antiprotozoal activities against *Acanthamoeba spp.* *J. Fluor. Chem.* *164*, 10-17.
- [6] Cichowicz, N. R., Nagorny, P. (2012) Synthesis of Conjugated Polyenes via Sequential Condensation of Sulfonylphosphonates and Aldehydes. *Org. Lett.* *14*, 1058-1061.
- [7] Rydzik, A. M., Leung, I. K. H., Kochan, G. T., Loik, N. D., Henry, L., McDonough, M. A., Claridge, T. D. W., Schofield, C. J. (2014) Comparison of the substrate selectivity and biochemical properties of human and bacterial gamma-butyrobetaine hydroxylase. *Org. Biomol. Chem.* *12*, 6354-6358.