Supporting Information for

Organophosphonate-degrading PhnZ reveals an emerging family of HD-domain mixed-valent diiron oxygenases

Bigna Wörsdörfer,^{1,*} Mahesh Lingaraju,² Neela Yennawar,³ Amie K. Boal,^{1,2} Carsten Krebs,^{1,2} J. Martin Bollinger, Jr.,^{1,2,*} Maria-Eirini Pandelia^{1,*}

¹Department of Chemistry, ²Department of Biochemistry and Molecular Biology, and ³Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA 16802, USA

*Corresponding authors:

mpandelia@gmail.com; jmb21@psu.edu; bigna.woe@gmail.com

Materials

Buffers, salts and other chemicals were purchased from commercial suppliers unless otherwise stated. Metallic ⁵⁷Fe was purchased from Isoflex (USA).

Protein production and purification

The genes for PhnZ (GeneBank ID: ACU83550) and PhnY (GeneBank ID: ACU83549) were codon-optimized for expression in *E. coli*, synthesized, and inserted into the plasmid pRSET_A_A185 at its *Xhol* and *EcoRl* restriction sites by GeneArt. The resulting plasmids pRSET_*phnZ* and pRSET_*phnY* encode the PhnZ and PhnY proteins with N-terminal His₆-tags. For PhnY, the codon-optimized gene was PCR amplified from the purchased pRSET_*phnY* plasmid and subcloned into a pET28a vector using *Ndel* and *Xhol* restriction enzymes, resulting in the plasmid pET28a *phnY* encoding the *phnY* gene with an N-terminal His₆-tag.

PhnZ was overexpressed in E. coli BL21 cells that had been transformed with the plasmid pRSET phnZ. Cultures were grown at 30 °C in enriched LB medium supplemented with ampicillin (0.15 mg/mL) to an OD₆₀₀ of 0.7, at which point PhnZ production was induced by addition of isopropyl-β-D-1-thiogalactopyranoside (IPTG) to 0.5 mM. 4 hours after induction, the cells were harvested by centrifugation at 6,000 g for 15 min at 4 °C. To obtain ⁵⁷Fe-labeled PhnZ for Mössbauer spectroscopy, E. coli BL21 cells containing the plasmid pRSET_phnZ were grown in M9 minimal medium (with 0.4% (w/v) glucose) supplemented with ampicillin (0.15 mg/mL) and 100 μ M ⁵⁷Fe^{II} at 37 °C to an OD₆₀₀ of 0.7, at which point PhnZ production was induced by addition of IPTG to 0.5 mM and the temperature was reduced to 18 °C. 20 to 24 hours after induction, the cells were harvested by centrifugation. The cell mass was resuspended in 4 mL/g lysis buffer (50 mM MOPS pH 7.5, 300 mM NaCl, 10 mM imidazole) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF), and the cells were lysed by two passages through a French Pressure cell. The cell lysate was cleared by centrifugation at 22,000 g for 30 min at 4 °C and loaded onto a gravity flow Ni²⁺-NTA agarose (Macherey-Nagel) column that had been equilibrated with lysis buffer. After the column was washed with 2 column volumes (CV) lysis buffer, 2 CV lysis buffer containing 20 mM imidazole, and 2 CV lysis buffer containing 40 mM imidazole, PhnZ was eluted with buffer containing 50 mM MOPS pH 7.5, 100 mM NaCl, and 250 mM imidazole.

For overexpression of PhnY, *E. coli* BL21 cells that had been transformed with the plasmid pET28a_*phnY* were grown at 30 °C in enriched LB medium supplemented with kanamycin (0.05 mg/mL) to an OD₆₀₀ of 0.7, at which point PhnY production was induced by addition of IPTG to 0.1 mM and the temperature was reduced to 15 °C. 25 hours after induction, the cells were

2

harvested by centrifugation at 6,000 *g* for 15 min at 4 °C. Cleared cell lysate was prepared as described for PhnZ and loaded onto a gravity flow Ni²⁺-NTA agarose (Macherey-Nagel) column that had been equilibrated with lysis buffer. After the column was washed with 2 CV lysis buffer and 2 CV lysis buffer containing 40 mM imidazole, PhnY was eluted with buffer containing 50 mM MOPS pH 7.5, 100 mM NaCl, and 250 mM imidazole.

The eluted proteins were concentrated in Centriprep YM-10 centrifugal concentrators (Millipore) and exchanged into MOPS buffer (50 mM MOPS pH 7.5, 10% (v/v) glycerol) using a PD-10 column (GE Healthcare). Protein purity was assessed by SDS-PAGE, and protein concentrations were determined using $\varepsilon_{280nm} = 31.4 \text{ mM}^{-1} \text{ cm}^{-1}$ for PhnZ and $\varepsilon_{280nm} = 48.36 \text{ mM}^{-1} \text{ cm}^{-1}$ for PhnY as determined on the basis of the amino acid sequence by the ProtParam tool of ExPASy (http://web.expasy.org/protparam/) (1). The Fe content was determined by ferrozine assay (2). Iron and other trace metals in the samples were quantified by inductively-coupled plasma atomic emission spectroscopy (ICP-AES) by Henry Gong at the Penn State Materials Characterization Laboratory (3). PhnZ typically contained 1.2 ± 0.2 Fe per monomer and no significant quantities of other divalent metals.

Enzymatic synthesis of 2-amino-1-hydroxyethylphosphonate by PhnY

The 200 mL reaction contained 5 μ M PhnY, 10 μ M Fe(NH₄)₂(SO₄)₂, 3 mM α -ketoglutarate (α -KG), 100 μ M sodium L-ascorbate, and 2 mM 2-aminoethylphosphonate (2-AEP) in 25 mM ammonium acetate buffer pH 7.5. Initially, PhnY was added to an O₂-free mixture of the Fe(NH₄)₂(SO₄)₂, L-ascorbate, α -KG, and 2-AEP in 26 mL ammonium acetate buffer after adjusting the pH to 7.5 with concentrated ammonium hydroxide (NH₄OH). The reaction was transferred outside the anoxic chamber, initiated by the addition of 174 mL air-saturated ammonium acetate buffer, and then continuously flushed with hydrated air under gentle stirring at room temperature. Conversion of 2-AEP to 2-amino-1-hydroxyethylphosphonate (OH-AEP) was monitored by ³¹P-NMR: 480 μ L aliquots were taken and quenched by heating to 85 °C for 15 min, and NMR samples were prepared as described in the ³¹P-NMR spectroscopy section below. After complete conversion of 2-AEP to OH-AEP, the reaction was quenched by incubation at 80 °C for 20 min, and the precipitate was removed by centrifugation at 7,500 × *g* for 15 min at 4 °C followed by filtration through a 0.2 μ M filter.

The product OH-AEP was purified using a gravity-flow column with 20 mL of Q-Sepharose FF resin (GE Healthcare). The pH of the filtered solution was adjusted to 8.5 with concentrated ammonium hydroxide and then loaded onto the Q-Sepharose column that had been equilibrated with 25 mM ammonium acetate buffer pH 8.5. OH-AEP was eluted from the column by 3 CV of

3

25 mM ammonium acetate buffer pH 8.5. The L-ascorbate and succinate from the reaction mixture remain bound to the column under these conditions and elute at ammonium acetate concentrations of > 200 mM. The eluted fractions were analyzed by mass spectrometry (ZQ 2000 mass spectrometer, Waters) and ³¹P-NMR spectroscopy, and fractions containing OH-AEP (the flow-through and the three 25 mM ammonium acetate elutions) were pooled and evaporated to dryness on a rotary evaporator. The solid was dissolved in a minimal volume of H₂O, and the concentration of OH-AEP was determined by ³¹P-NMR spectroscopy using 2-AEP as an internal standard. The OH-AEP solution was aliquoted and dried using a speedvac concentrator (Savant SPD131DDA, Thermo Scientific).

The specific optical rotation of the purified OH-AEP dissolved in H₂O was determined to be $[\alpha]_{D}^{20} = -40.2$ in good agreement with the previously determined literature value for *(R)*-OH-AEP (-31.4) (4).

Activity assays for PhnZ

For activity assays, the Fe^{II}/Fe^{III} state of the protein was generated by reduction of the purified protein with L-ascorbate; the fully oxidized Fe^{III}/Fe^{III} state was prepared by treatment of the protein with potassium ferricyanide (K₃[Fe(CN)₆]). 200 μ M PhnZ was incubated in the absence of O₂ with 5 equiv. L-ascorbate (relative to the diiron cofactor), 0.2 equiv. ferricyanide, or 0.2 equiv. ferricyanide followed by 5.2 equiv. L-ascorbate in 50 mM MOPS pH 7.5, 10% (v/v) glycerol for a several minutes. EPR samples were taken to determine the yield of the mixed-valent Fe^{II}/Fe^{III} cofactor in these preparations. For the reaction in the absence of O₂, PhnZ was diluted with deoxygenated MOPS buffer to a final concentration of 10 μ M, and OH-AEP substrate was added to 2 mM. The solution was incubated anaerobically under gentle stirring at room temperature. For reactions were continuously flushed with hydrated air under gentle stirring at room temperature.

At various time points, 480 μ L aliquots were taken and quenched by addition of 4% (v/v) concentrated acetic acid. Conversion of OH-AEP to phosphate was detected by ³¹P-NMR, as described in the ³¹P-NMR spectroscopy section below.

PhnZ activity assays for testing O₂ incorporation into the product

Reactions contained 10 μ M PhnZ enriched in the Fe^{II}/Fe^{III} state, 1 mM OH-AEP and were initiated by the addition of ¹⁸O₂- or ¹⁶O₂-saturated H₂O (~1.8 mM O₂), resulting in a final O₂

concentration of ~1.6 mM. After 1 h the reactions were stopped by addition of 0.5% (v/v) concentrated acetic acid. Precipitation was removed by centrifugation. Incorporation of oxygen from O_2 into the product glycine was detected by LC-MS, as described in the LC-MS section below.

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

LC-MS experiments were carried out on an Agilent 1200 series LC system coupled to a triple quadrupole mass spectrometer (Agilent 6410 QQQ LC/MS; Agilent Technologies). 60% (v/v) acetonitrile was added to the samples and they were filtered (0.2 µm filter) before injection on a ZIC-HILIC (Merck) column that had been equilibrated with 5% solvent A (95% H₂O, 5% acetonitrile, 0.1% formic acid) and 95% solvent B (5% H₂O, 95% acetonitrile, 0.1% formic acid). The reaction mixture was separated at 0.3 mL/min by applying a linear gradient of 5-20% A over 15 min, isocratic at 20% A for 5 min, a linear gradient of 20-70% A over 15 min, then returning to 5% A over 15 min and re-equilibrating at 5% A for 5 min. Detection of the glycine products was performed using electrospray ionization in the positive mode (ESI+) with single ion monitoring at *m*/*z* of 76.1 for detection of ¹⁶O-glycine (H₃N-CH₂-C¹⁶O¹⁶OH)¹⁺ and *m*/*z* of 78.1 for detection of ¹⁸O-glycine (H₃N-CH₂-C¹⁶O¹⁸OH)¹⁺. Relative quantities of ¹⁶O-glycine and ¹⁸O-glycine were determined by integrating the corresponding peak using the MassHunter QualitativeAnalysis Software (Agilent Technologies).

Rapid freeze-quench EPR sample preparation

An anaerobic solution of PhnZ containing 0.18 mM Fe^{II}/Fe^{III} cofactor and 3 mM (*R*)-OH-AEP was prepared and transferred to a gas-tight freeze-quench (FQ) syringe. A second syringe was filled with O₂-saturated buffer (~1.8 mM O₂, 50 mM MOPS pH 7.5, 10% (v/v) glycerol). The solution of the substrate-bound complex (PhnZ-Fe^{II}/Fe^{III}•(*R*)-OH-AEP) was mixed at 5 °C with 0.5 equivalent volumes of the O₂-saturated buffer in the FQ set-up that has previously been described (5), affording final concentrations of 0.121 mM PhnZ-Fe^{II}/Fe^{III}, 2 mM (*R*)-OH-AEP and 0.6 mM O₂. The reaction solution was aged for the desired time and then ejected into the isopentane cryo-solvent (~ -150° C) to terminate the reaction.

³¹P-NMR spectroscopy

For ³¹P-NMR measurements, sodium dithionite, EDTA and D₂O were added to the samples to final concentrations of 0.2 mM, 0.2 mM, and 20% (v/v), respectively. Solution ³¹P-NMR spectra of the various phosphorous containing compounds were recorded at room temperature

in an AVX-360 Bruker spectrometer. The spectra were recorded using a 1D sequence with power-gated ¹H decoupling. Chemical shifts are quoted with respect to an 85% phosphoric acid solution at 0 ppm. The 85% phosphoric acid solution was transferred into a 2 mm OD quartz tube, which was then inserted into the 5 mm OD quartz NMR tube containing the solutions under investigation. NMR spectra were further processed with the freely available Spinworks (version 1.3.8.1) software (Dr. Kirk Marat, University of Manitoba, Canada).

Mössbauer spectroscopy

Mössbauer spectra were recorded on spectrometers from WEB Research (Edina, MN). The spectrometer used to acquire the weak-field spectra is equipped with a Janis SVT-400 variable-temperature cryostat. The spectrometer used to acquire the strong-field spectra is equipped with a Janis 8TMOSS-OM-12SVT variable-temperature cryostat. The external magnetic field was applied parallel to the γ -beam. All isomer shifts quoted are relative to the centroid of the spectrum of α -iron metal at room temperature. Simulation of the Mössbauer spectra was carried out by using the WMOSS spectral analysis software (www.wmoss.org, WEB Research, Edina, MN). The high-field simulations were carried out using the spin-Hamiltonian (6), in which all symbols have their usual meaning.

$$H = \mu_B \mathbf{S} \cdot \mathbf{g} \cdot \mathbf{B} + D \left(\mathbf{S}_{\mathbf{z}}^2 - \frac{S(S+1)}{3} \right) + E \left(\mathbf{S}_{\mathbf{x}}^2 - \mathbf{S}_{\mathbf{y}}^2 \right) + \sum_{i=1}^2 \frac{e Q V_{zz,i}}{4} \left[\mathbf{I}_{z,i}^2 - \frac{I_i (I_i+1)}{3} + \frac{\eta}{3} \left(\mathbf{I}_{x,i}^2 - \mathbf{I}_{y,i}^2 \right) \right] + \sum_{i=1}^2 \mathbf{S} \cdot \mathbf{A}_i \cdot \mathbf{I}_i - \sum_{i=1}^2 g_n \mu_n \mathbf{B} \cdot \mathbf{I}_i$$

The first term corresponds to the electronic Zeeman interaction, the second and third correspond to the axial and rhombic zero field splitting, the fourth term describes the interaction between the electric field gradient and the nuclear quadrupole moment, the fifth term corresponds to the hyperfine interactions between the electronic spin and the ⁵⁷Fe nucleus, and the sixth term represents the nuclear Zeeman interaction. The simulations were carried out with respect to the total spin of the cluster.

EPR spectroscopy

EPR spectra were acquired on a Bruker ESP300 CW X-Band spectrometer (operating at approx. 9.48 GHz) equipped with a rectangular cavity (TE_{102}) and a continuous-flow cryostat (Oxford 910) with a temperature controller (Oxford ITC 503).

General crystallographic methods

Preliminary screening of protein crystals was performed at the X-ray crystallography facility in the Huck Institutes for the Life Sciences. Diffraction datasets for structure solution were collected at the Life Sciences Collaborative Access Team (LS-CAT) beamlines at the Advanced Photon Source. All datasets were indexed, integrated, and scaled using the HKL 2000 software package (7). Model building was accomplished in Coot (8), and refinement was carried out in PHENIX (9). All data collection and refinement statistics are reported in Table S1. Structure validation was performed with Molprobity. Figures were generated with the PyMOL Molecular Graphics System (Schrödinger, LLC), and electron density maps were calculated with PHENIX (9).

Determination of the structure of citrate-bound PhnZ

N-terminally His₆-tagged PhnZ overexpressed from pRSET_phnZ was prepared for crystallographic analysis as described in the protein production and purification section. An additional size-exclusion chromatography step (Sephacryl-200 gel filtration column, GE Healthcare) was included using 50 mM MOPS pH 7.5, 10% (v/v) glycerol as running buffer. Hexagonal prism-shaped crystals of PhnZ (12 mg/mL) appeared after one week of incubation at room temperature using the sitting drop vapor diffusion method with 0.1 M HEPES pH 7.5, 1.4 M trisodium citrate as the precipitant. Crystals were mounted on rayon loops for data collection and flash-frozen by direct plunge into liquid nitrogen with no further cryoprotection. The structure was solved by single-wavelength anomalous diffraction (SAD) phasing using a dataset collected at the Fe X-ray absorption peak. The phases were determined with autoSHARP (10) using the two active site Fe ions (Table S1). From the resulting electron density maps, an initial model was generated with Buccaneer (11). An additional high-resolution native dataset was used to generate and refine the final model. The final model consists of PhnZ residues 1-66, residues 72-187, two Fe ions, one citrate anion, one bridging hydroxo group and 153 water molecules. No electron density is observed for the N-terminal affinity tag and for a surface loop (residues 66-72) near the active site. The occupancies of two metal ions were refined with PHENIX and both sites are modeled at nearly full (92%) occupancy. Ramachandran plots show that 100% of the residues are in allowed and additionally allowed regions.

Determination of the structure of OH-AEP-bound PhnZ

The *phnZ* gene was PCR amplified from the pRSET_*phnZ* plasmid and subcloned into a pET26b vector using *NdeI* and *XhoI* restriction sites resulting in the plasmid pET26b_*phnZ*

encoding the *phnZ* gene with a C-terminal His₆-tag. This protein was overexpressed in *E. coli* BL21 cells that had been transformed with pET26b_*phnZ*. Cultures were grown at 30 °C in enriched LB medium supplemented with kanamycin (0.05 mg/mL) to an OD₆₀₀ of 0.7, followed by IPTG induction (0.1 mM) for 16 hours at 30 °C. The protein was purified and prepared for crystallization trials as described for the N-terminally tagged protein overexpressed from pRSET_*phnZ*.

Crystals of C-terminally His₆-tagged PhnZ with bound (*R*)-OH-AEP were prepared from a solution containing both the enzyme (14 mg/mL) and the substrate (8 molar equiv.) that was preincubated for 20 minutes on ice prior to setup of the crystallization trials. Plate-like crystals grew overnight in sitting drop vapor diffusion trays with 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate, and 25% (w/v) PEG 4000 as the precipitant. Crystals were mounted on rayon loops for data collection and incrementally transferred (5% increase in glycerol) into a final cryoprotection solution of 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate, 25% (w/v) PEG 4000 and 25% (v/v) glycerol.

The structure was solved by molecular replacement using PHASER (12) with the citratebound PhnZ structure as the search model. Four substrate-bound PhnZ molecules are present in the asymmetric unit. The coordinates for (*R*)-OH-AEP were generated using the PRODRG server (13). The final model consists of residues 4-65 and residues 70-187 in chains A, B, and E; residues 4-64 and residues 70-187 in chain D; eight Fe ions; and four (*R*)-OH-AEP molecules. No electron density was observed in any of the four PhnZ molecules for residues 1-3, residues 66-69, or the C-terminal His₆ tag. The occupancies of the substrate atoms and metal ions were refined in PHENIX and are modeled at 100% occupancy in the final model. Ramachandran plots show that 99.8% of the residues are in allowed and additionally regions.

Phylogenetic analysis

The Basic Local Alignment Search Tool (BLAST) was used to perform similarity searches through the NCBI BLAST webpage interface (http://www.ncbi.nlm.nih.gov/BLAST/) (14). The retrieved sequences were initially grouped on the basis of similarity relationships and the genomic contexts of their encoding genes. The sequences were further culled by requiring that each contain at least the six conserved residues necessary for coordinating a dinuclear cluster (i.e. four histidines and two glutamates). Each selected sequence was examined for completeness and length. Approximately 8 sequences from each group were chosen, so as to have equal representations from each group within the phylogenetic tree and avoid bias. The amino acid sequences of each group were aligned using the MAFFT (version is 7.03,

8

http://mafft.cbrc.jp/alignment/software/) online webserver (15) and the ClustalW algorithm as implemented in the MEGA 5.2 software (16, 17). For the construction of the tree, all sequences were further re-aligned together. Neighbor joining (NJ) (18), minimum evolution and maximum likelihood (19) trees were constructed using the respective algorithms in the MEGA software package and the MAFFT online server as well as the older software ClustalX2 (20). Similar trees were obtained with the different methods. The robustness of the trees was further confirmed by using the bootstrap method with 1000 replications.



Figure S1. 4.2-K Mössbauer spectrum of the aerobically isolated PhnZ recorded in a wide range (~ ±12 mm/s) of Doppler velocities. An external magnetic field of 8 T applied parallel to the γ -beam. The orange solid line represents a simulation of the mononuclear high-spin Fe^{III} contaminant(s). Its integrated intensity corresponds to 7% of the total Fe species in the sample. The simulation was carried out using the spin-Hamiltonian formalism with the following parameters: S = 5/2, D = -1.0 cm⁻¹, E/D = 0.33, g = 2.0, $\delta = 0.55$ mm/s, $\Delta E_Q = -1$ mm/s, $\eta = -2.5$, $A_{iso}/(g_N \beta_N) = -22.2$ T.



Figure S2. 4.2-K/0-T Mössbauer spectrum of the dithionite-reduced form of PhnZ. The experimental spectrum (black vertical bars) is fit by two sets of quadrupole doublets with parameters $\delta_1 = 1.30$ mm/s, $\Delta E_{Q1} = 3.43$ mm/s (blue solid line), $\delta_2 = 1.3$ mm/s and $\Delta E_{Q2} = 2.83$ mm/s (green solid line) and a Lorentzian line shape. The overall fit (summation) is shown as a red solid line.



Figure S3. Comparison of the overall structures of citrate-bound PhnZ (A) and *M. musculus* MIOX (PDB accession code 2HUO) (B). A disordered region in PhnZ is represented by a black dashed line. The Fe ions in both structures are shown as orange spheres, and the substrate and citrate molecules are shown in stick format.

PhnZ-uncultured	$\begin{array}{ccc} \alpha 1 & \eta 1 & \alpha 2 \\ \alpha $	2
PhnZ-uncultured P.marinus MIT 9301 P.marinus MIT 9303 Frankia.sp M.colombiense P.pacifica P.maris A.niger T.stipitatus P.chrysogenum consensus	I IQU ZQ 30 40 	<pre></pre>
PhnZ-uncultured	α_3 η_2	α4 202020202
PhnZ-uncultured P.marinus MIT 9301 P.marinus MIT 9303 Frankia.sp M.colombiense P.pacifica P.maris A.niger T.stipitatus P.chrysogenum consensus	50 60 70 80	90 /GADYLRGLGFS LASNFLK.DFFS /AARALS.EFFP /GARYLT.RWFG /AASILG.RWFP IGADWLDGLGFD /GAQULA.PFFG IGAEYLRSLGFS IGAHYLQSLGFS IGASYLRSLGFS . * *
PhnZ-uncultured	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	α9 2020202020 160
PhnZ-uncultured P.marinus MIT 9301 P.marinus MIT 9303 Frankia.sp M.colombiense P.pacifica P.maris A.niger T.stipitatus P.chrysogenum consensus	ERVACLIEGEVAAKRYLVSSKASYLKNLSDASRATLEYOGGPMDEGERRLFEEREDFKU EEITESIRLHVVAKRYLCSIDNSYYESLSKASKNSFKIQGGALNKEEINELENNKYFKU SNVVTPVQRHVSAKRLLCSLDGAYYEGLSEASKRSFAVQGGPMSFEAERLQALEGMQU PDIGTPVALHVKAKRYLVATDPAYANSLSVASTRSLRVQGGPMTGREVDAFLCLPHADI PAVTAPIALHVAAKRYLVATDPSYACGLSPASVQSLRVQGGPMNAEEIVHFGRLRHAGC PRVGELVRAHVDAKRYLVATPPAYAARLSDASRQTLALQGGPMDAEUVHFGRLRHAGC KAVTEPIFLHVPAKRYLCSVNADYFNGLSRASQRSVELQGGLMTDAEIAEFEQNPFHHT EKVCRLVESHVPAKRYLTATNKAYYETLSSASKKSLEFQGGPFKGAELEAFETDPLRDE ETVCRLVKSHVAAKRYLTATNKSYYNSLSSASQRSLAFQGGPFLGNELEAFEGDPLRDE EVVCRLVNSHVAAKRYLTATDQRYHGSLSSASQKSLAFQGGPFRDADLRNFEEDPLRDE	CLKIRAWDEKG AVRLRKWDDRG ANALRRWDDRA AIALRRWDDRA AKLRVRSWDERA CAVLLRRWDDGA SMVKLLWDDAA SMVSLRLWDDAA SMVSLRLWDDAA
PhnZ-uncultured	$\begin{array}{ccc} \eta 4 & \alpha 10 \\ 2 \cdot \cdot 20202020000 \\ 170 & 180 & 190 \end{array}$	
PhnZ-uncultured P.marinus MIT 9301 P.marinus MIT 9303 Frankia M.colombiense P.pacifica P.maris A.niger T.stipitatus P.chrysogenum Consensus	KQTDLKVPGPEHYRKMMEEHLSENQNImage: Construct of the system	on sidue

Figure S4. Multiple sequence alignment of PhnZ-like sequences.



Figure S5. Stereoview of the $2F_{o}$ - F_{c} electron density map (gray mesh contoured at 2.0 σ) for the diiron site and bound citrate molecule in the high-resolution PhnZ X-ray structure. The Fe ions are shown as orange spheres and the bridging oxygen species is shown as a red sphere. The side chains of the coordinating residues are shown in stick format and colored by atom type.



Figure S6. Stereoview of an omit map (green mesh contoured at 3.0σ) showing the electron density for the substrate molecule in chain A of the substrate-bound PhnZ structure. The Fe ions are shown as orange spheres and the side chains of the coordinating residues are shown in stick format and colored by atom type. The substrate molecule was modeled at 100% occupancy in all monomers with minor differences in the orientation of the amino group.



Figure S7. ³¹P-NMR spectrum of a sample containing 10 μ M PhnZ, 2 mM OH-AEP and 5 equiv. of L-ascorbate that was incubated in the anoxic chamber for 6 hours at ambient temperature. The absence of phosphate product establishes that O₂ is required for the PhnZ reaction.



Figure S8. Incorporation of an O-atom from O₂ into glycine demonstrated by LC-MS single-ionmonitoring (SIM) chromatograms of PhnZ reactions in the presence of ¹⁶O₂ and ¹⁸O₂ and of a no enzyme control. SIM chromatograms for m/z = 76, corresponding to ¹⁶O-glycine (H₃N-CH₂-C¹⁶O¹⁶OH)¹⁺, are shown in black, and those for m/z = 78, corresponding to ¹⁸O-glycine (H₃N-CH₂-C¹⁶O¹⁸OH)¹⁺ are shown in red. The nearly complete (98%) incorporation of ¹⁸O from ¹⁸O₂ into the glycine product establishes that PhnZ is an oxygenase.



Figure S9. 120-K/0-T Mössbauer spectra of PhnZ. (A) Spectrum of the aerobically isolated PhnZ following treatment in the absence of O_2 with excess L-ascorbate, which was then removed using a PD-10 desalting column (black). The spectrum contains an upper limit of 4% of the Fe^{II}/Fe^{II} state (red) and approximately 46% of the Fe^{III}/Fe^{III} form (blue). Subtraction of these contributions yields the reference spectrum of the mixed-valent Fe^{II}/Fe^{III} form (B; black). The individual contributions of the Fe^{III} site (δ = 0.51 mm/s, ΔE_Q =1.03 mm/s) and the Fe^{III} site (δ = 1.20 mm/s, ΔE_Q = 2.55 mm/s) are shown in green and orange, respectively. The summation of these contributions is shown in purple.



Figure S10. Reaction of PhnZ-Fe^{II}/Fe^{III}•OH-AEP with O₂ at 5 °C monitored by stopped-flow absorption spectroscopy. (Left) Absorption spectra at the reaction times shown in the figure. (Right) Kinetics of the absorbance at 406 nm. Concentrations after mixing were 0.07 mM PhnZ-Fe^{II}/Fe^{III}, 1.5 mM OH-AEP, and ~ 0.9 mM O₂ in 50 mM MOPS buffer pH 7.5.

		T				ΤΤ			
B.bacteriovorus	тт	0000000	α6 20000000	.000	α7	0000	η3 000		η4 000 00
14	4 <u>0</u>	150	160	17	7 0 1	80*	190		200
B.bacteriovorus B.marinus	IENTDK	TISHHGVT	ZSTLSIAL ZSALATKI	. AQKLGITI . GOKMKLTE	DPKKTQLLTL S. EOLDETAV	GALLHD AGLLHD	YGHHHSPI TGLMOLDI	LNLNQP EKTRELENKI	. LDSMSPE JKKDFTPE
B.exovorus	LTGHDF	YTYNHSFD	7SIYSLGL	.GQVLGMDH	K.VTLEELGQ	SSLFHD	IGKRNVP	LEILCK	.KGALSED
R.balticaSH1 R balticaWH47	VLNHDY	ATFTHSAN	ALYAGML	ASELGMT(2. REVELVVA	GGLLHD	LGKLEIPI		PGRLDED
GeobacterM18	LVDHNA	YTYVHSVQ	ATYSISL	HVKMFELSI	R.DELMDVG <mark>I</mark>	GSLFHD	YGKVYVPI	RALLDK	. PGKLSPA
D.aespoeensis	FISHDY	KTYTHCMH	FIYSVAV	. FQAFDMSH	E.ADTYECGL	GALLHD	IGKARIPI	RRILNK	RGSLTQA
D.fructosovoran V.cholerae	AKSO	FTFOHSOK	AELCOHL	. LQSMKADI . AKELGLNV	S.YTQVQAGI /.EMOKALYL	TGLVHD	IGKLHTP	GEILSK EEILHK	. PGPLTPE
V.choleraeLMA39	AKSQ	FTFQ <mark>H</mark> SQK	/AELCQH <mark>L</mark>	AKELGLNV	/.EMQKALY <mark>L</mark>	TGLVHD	I <mark>G</mark> KLH <mark>TP</mark> I	EE <mark>IL</mark> HK	. PGK <mark>L</mark> NES
V.mimicus V.splendidus	AKSQ	FTFYHSQK	AEICQRL	AKELGLDA	A.EMQKALYL I TTORKLYL	TGLVHD	IGKLHTPI	EKILHK	PDKLNES
ShewanellaANA-3	AKSP	FTFHHSDK	ALLAKLV	.AKDCGISI	D. TDAELLYV	AGLLHD	VGKLKTPI	DLLLHK	EGKLTRE
AzoarcusBH72	AKSP	YTDDHSRR	AAISRHL	. AGALGRDH	R.DTLEMIEL	AGLLHD	MGKLRVPI	DEIIEK	. PGALTRE
M.magneticum C.thalassium	AKSP	FTVAHSIG	AALAAQI	. AGDMGLDA . AKLCGLSA	A.ETVRLVEI K.ENSKKIEA	AGYLHD	IGKLHVPI	DEILEK	. PGPLNAO
	~ ~ ~ ~								~
		↓		0	$\downarrow \downarrow$		-	↓.	0
B.bacteriovorus	000000	0.8	00 TT	α9 0000000	00	тт	η5	1x 0 0 0 0 0 0 0 0 0	0000000
2.240000145	210	22	20	230	240	25	۰ ۰	260	270
B.bacteriovorus	DLALWK	KHPIEGAQI	KVQDKKHFI	DQTVINIIC	GQHEETINGT	G.PKGL	REKDMDPI	LAVLVSSANA	AMDRLIT.
B.marinus B.exovorus	EKLAYG	OHPAYGLK	LKDKPYVI LNAHENI (SDAVKAACY	CNHEEVKSGH ZEHHESWTGG	G.PN GYPODI	SGHEIHPI	SEYILSLVN. FGRIVALTD	LYDKTII. FYDAMTTO
R.balticaSH1	EFALIK	K <mark>H</mark> PGDGFK(QLALRDDL'	ſFG <mark>Q</mark> LMMVY	Q <mark>H</mark> H <mark>E</mark> RID <mark>G</mark> G	GYPVGT	VGEDIHPU	WGRLCAVVD	IYEAVTSQ
R.balticaWH47_	EFALIK	KHPGDGFK(LALRDDL'	FFGQLMMVY	QHHERIDGG	GYPVGT	VGEDIHPU	WGRLCAVVD	I YEAVTSQ
D.aespoeensis	ERDVIK	EHPVHGVSN	ACAHLPM.	TQNTINCII	FHHEKLDGT	GYPAGL	RRESIPLI	PVRIISVSD	IYDALTTA
D.fructosovoran	ERAVIN	THPAKGLAI	LCQDVPL.	IQT <mark>A</mark> AHCII	MHHERMDGA	GYPGGL	SGELIPP	YVRALAVAD	VYDALTTK
V.cholerae V.choleraeLMA39	EYLCIQ	RHSTDSRY	LLQMVFG.(LLOMVFG.(2SVVCEWAC SVVCEWAC	GNHHERLDGS GNHHERLDGS	GYPRGL GYPRGL	QGAAIDLI OGSAIDLI	PSRIIAIAD	/FQALTQA /FOALTOA
V.mimicus	EYLCVQ	RHSTDSRY	LQMVFG.	KSILGEWAC	GN <mark>H</mark> H <mark>E</mark> RLD <mark>G</mark> S	GYPRGL	QGSEIDLI	PSRIIAVAD	VFQALTQA
V.splendidus	EYCCIK	RHATDTRFA	ALQELFS.	SPQVCQWAS	SDHHERLDGS	GYPMGK	TAEDLDQI	PSRIVAVVD	/FQALTQS
AzoarcusBH72	EKACIS	RHSYDTYR	LARIFP.	STPIPDWAG	GAHHENLLGQ	GYPFRR	RAGEIDLI	ETRIISVAD	FOALLOD
M.magneticum	ERVLMK	RHAFDTYDY	/LFRLFG.I	DNPIAQWAA	AF <mark>H</mark> H <mark>E</mark> TLN <mark>G</mark> E	GY <mark>P</mark> FQL	SAQALPLI	PARIIAVSD	IFQALVQE
C.thalassium	ERAIIE	QHSFETYE	LLRRISG.	I E D <mark>I</mark> A K W A A	AFHHETVSCN	GYPFRA	нтеецот	EARIIAVAD	<u>LFQALAQ</u> D
		α11			α12				
B.bacteriovorus	••2	280	290	200	200000000				
B.bacteriovorus	FEGV	PKAEAAKK	LMIDHV.	.GKHPLQHI	QHLNDILKG				L
B.marinus	<mark>T</mark> TD	LTPK <mark>EA</mark> IKI	ILIDEL.	.GNFELÕLI	ĴÑKFKEVLTE			EG.	LLDL
B.exovorus R balticaSH1	RSYNTP	LKPTEALN	MMKDKIS	. GRFDPDMI	LNAMYSILFK			MEK.	DQAS
R.balticaWH47	RPYRTP	MSREDACN	LIRRESG	. KALDPEM	/ECWISIIHS				TMPK
GeobacterM18	RSYRQA	LSKESALE	IMYGEMD	.GSFDMQYI	LNTFRDTLH.				
D.aespoeensis D.fructosovoran	RPYAEA	MPPYEALT. MOPFOALR	LIRHQMR LMRDEMG	. DGVDMDIH . GAFDLDVH	KKFVAVLSG KRLVVII.SG				ADMI
V.cholerae	RPYRGS	MSLNEVMN	IMRHEVS	CGRLDSQVI	DVIVRNSQQ	YYQLSI	AESPTEWA	A	
V.choleraeLMA39	RPYRGS	MSLNEVMN	IMRHEVS	CGRLDSQVE	DVIVRNSRQ	YYQLSI	AESPTEWA	A	
V.splendidus	RPIRGR	MTLEETLA	ILTEOVDS	SYKLDREVI	FECLRKHAEY	CFELST	DKRM	· · · · · · · · · · ·	YAF
ShewanellaANA-3	R P Y <mark>R</mark> G S	MSLS <mark>EV</mark> LE	IMLPMVQ	QGKLDSNVY	DVLLADADN	FYQLST	QE		YEA
AzoarcusBH72	RPYRAR	LSGEEVMQ	RIEALVDI	EGRLDAAV	SALRREFDT	CVHLAS	GADA	LRDHFI	PELTA
C.thalassium	RPYREG	LALKEILR	IMNALVE	KGNLDESLV	JALVNENALA	CYHEAT	RTDI	· · · · · · · · · · · ·	- GRIVRI

Figure S11. Multiple sequence alignment of the proteins in the HD-GYP and HD-[HD-GYP] clades starting from residue 140. Residues that coordinate the metal ions are indicated with purple arrows.

			2	-2	
B.subtilis	i 10		α2 0000000000000 3 0 *	<u>0000000000</u> 4 0 5 0	Q
B.subtilis B.amyl.N27 B.amyl.DSM7 B.amyl.DSM7 B.pumilus B.nethanolicus G.kaustophilus GeobacilusWCH7 B.licheniformis B.halodurans B.halodurans B.sbevis B.halodurans B.sibiricum E.sibiricum E.sibiricum E.gallinarum T.halophilus M.plutonius CarnobacteriumA L.saerimmeri S.agalactiae S.agalactiae S.salivarius L.buchneri S.termitidis	MNREEALACVK MNREEALACVK MNREEALACVK MNREEALACVK MNREEALACVK MNREEALACVK MNREEALACVK MNLDTALKIVK MREEALACVK MREEALACVK MRECALLIVK MRECALLIVK MRECALLIVK MREEALACVK MRERTALNKVK MREEALACVK MREALACVK	QQLTEHRYIHTI EQUITEHRYIHTI EQUITEHRYIHTI EQUITEHRYIHTI EQUITEHRYIHTI EQUITEHRYYTHTI EQUITEHRYYTHTI EQUITENRYYTHTI EQUITEQRYYQHTI HMMSDKRFNHVL SVMSSEKRFNHVL SVMSSEKRFNHVL ANVSSEKFFNHVL DAUTDSRYQHVL ARLDEKRYSHTK	GVMNTAIELAERFG GVTDTAVILAGRFG GVTDTAVILAGRFG GVTDTAVILAGRFG GVTDTAVILAGRFG GVTETAVALAERFG GVVETAVALAERFG GVVETAILEAERYG GVWETAIDLAKLYG GVVETAIQLAERFG GVVETAIQLAERFG GVETAIQLAERFG GVEAAALSLAERFG GVEAALSLAERFG GVEQAAILSLAERYG GVEQAAILSLAERYG GVEQAAISLAEKYG GVEQAAIELAERYG GVEQAAIELAERYG GVEQAAIELAERYG GVEQAAIELAERYG GVEQAAIELAERYG GVECAAIELAERYG GVECAAIELAERYG GVECAAIELAERYG GVECAAIALA	$ \begin{array}{l} \textbf{A} \textbf{D} \textbf{S} \textbf{K} \left[\textbf{K} \textbf{A} \textbf{E} \mid \textbf{A} \textbf{A} \textbf{I} \mid \textbf{F} \textbf{D} \right] \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{P} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{M} \textbf{A} \textbf{I} \textbf{F} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{F} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{M} \textbf{A} \textbf{I} \textbf{F} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{F} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{M} \textbf{A} \textbf{I} \textbf{F} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{F} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{M} \textbf{A} \textbf{I} \textbf{F} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{F} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{I} \textbf{A} \textbf{I} \textbf{F} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{L} \textbf{A} \textbf{A} \textbf{F} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{K} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{L} \textbf{A} \textbf{A} \textbf{F} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{K} \textbf{K} \textbf{K} \textbf{E} \textbf{L} \textbf{A} \textbf{A} \textbf{F} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{C} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{L} \textbf{A} \textbf{I} \textbf{F} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{D} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{L} \textbf{A} \textbf{I} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{C} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{L} \textbf{A} \textbf{I} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{D} \textbf{C} \textbf{K} \textbf{K} \textbf{A} \textbf{E} \textbf{L} \textbf{A} \textbf{I} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{A} \textbf{S} \textbf{C} \textbf{K} \textbf{K} \textbf{A} \textbf{S} \textbf{I} \textbf{A} \textbf{L} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{S} \textbf{S} \textbf{K} \textbf{K} \textbf{A} \textbf{S} \textbf{I} \textbf{A} \textbf{L} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{S} \textbf{S} \textbf{K} \textbf{K} \textbf{A} \textbf{S} \textbf{I} \textbf{A} \textbf{L} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{S} \textbf{S} \textbf{K} \textbf{K} \textbf{A} \textbf{S} \textbf{I} \textbf{A} \textbf{L} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{S} \textbf{S} \textbf{K} \textbf{K} \textbf{A} \textbf{S} \textbf{I} \textbf{A} \textbf{L} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{S} \textbf{S} \textbf{K} \textbf{K} \textbf{A} \textbf{S} \textbf{I} \textbf{A} \textbf{L} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{S} \textbf{S} \textbf{K} \textbf{K} \textbf{A} \textbf{S} \textbf{I} \textbf{A} \textbf{L} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{S} \textbf{S} \textbf{K} \textbf{K} \textbf{A} \textbf{S} \textbf{I} \textbf{A} \textbf{L} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{S} \textbf{V} \textbf{D} \textbf{E} \textbf{K} \textbf{K} \textbf{S} \textbf{S} \textbf{I} \textbf{A} \textbf{L} \textbf{H} \textbf{D} \textbf{Y} \textbf{A} \\ \textbf{S} \textbf{S} \textbf{K} \textbf{K} \textbf{S} \textbf{S} \textbf{I} \textbf{A} \textbf{S} \textbf{L} \textbf{M} \textbf{M} \textbf{M} \textbf{M} \textbf{A} \textbf{A} \textbf{S} \textbf{L} \textbf{D} \textbf{S} \textbf{M} \textbf{M} \textbf{S} \textbf{M} \textbf{S} \textbf{M} \textbf{M} \textbf{S} \textbf{M} \textbf{M} \textbf{S} \textbf{M} \textbf{S} \textbf{M} \textbf{S} \textbf{M} \textbf{S} \textbf{M} \textbf{S} \textbf{M} \textbf{M} \textbf{M} \textbf{M} \textbf{M} \textbf{M} \textbf{M} M$	K FREPPKKK FREPPKKKK FREPPKKKK FREPPKKKK FREPPKKKK FREPVKKK FREPVK FREPVKK FREPVKK FREPVKKKK FREPVKKK FREPVKKKK FREPVKKK FREPVKKKK FREPVKKKK FREPVKKKK FREPVKKKK FREPVKKKK FREPVKKKK FREPVKKKKK FREPVKKKKK FREPVKKKKKKK FREPVKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
B.subtilis	α4 η2 η3 <u>00000000000</u> <u>00000</u> <u>0000</u> 60 70 <u>80</u>	α5 200000000 9 9 *	α6 <u>00000000</u> 100 110	α7 TT <u>000000000</u> 129	<u>20</u> 130
B. subtilis B. anyl. M27 B. anyl. DSM7 B. anyl. TT-45 B. pumilus B. methanolicus G. kaustophilus Geobacilus%CH7 B. licheniformis B. halodurans B. pseudofirmus E. sibiricum E. sibiricum E. sibiricum E. sallinarum T. halophilus M. plutonius CarnobacteriumA L. saerimmeri S. agalactiae S. agalactiae S. agalactiae S. salivarius L. buchneri S. termitidis	EEMKQTIAREKMPAHLLDH.NPELWHAPY EBMKQIIVREKMPAELLDY.NAELWHAPY EEMKQIIVREKMPAELLDY.NAELWHAPY EEMKQIIVREKMPAELLDY.NAELWHAPY DEMKGIIQQEQMPPLLEY.SPELWHAPY DEMKGIILQQEQMPPLLEY.SPELWHAPY DEMKGIILQQEQMPPLLEY.NSELWHAPY EEMKQIILEQNMPNDLLAY.NSELWHAPY EEMKQIILEQNMPNDLLAY.NSELWHAPY NEMRTLIREKLSQQULLY.GDELLHAP QEMRDLIQSTLKEKDVLGY.GDELLHAP SEMQVIDESTLESUNG EEKQUILEQNMPSTELLG.EKELWHAPY DEFRQVIRDYDYLLEY.GDELLHAP EFRQVIRDYDYLLEY.GDELLHAP EFRQVIRDYDYLLEY.GDELLHAP EFRQVIRDYDYLLEUNS.NNAIWHGI SEFQUVIRDYDYLLLY.GNFIWHGAY SEFQLVIRDYDYLLLY.GNFIWHGAY DDFRKIIKQENLSPELLNY.GNFIWHGY DEFLRLIDKYQLDPELKKW.GNNVWHGIY QEFLLIDKYQLDPELKKW.GNNVWHGY CEFLLIDKYQLDPELKKW.GNNVWHGYY CDFIKAIHEYQLNPILLNY.GSIWHGYY	VGAYLVQREAGV VGAYLAEKEAGI VGAYLAEKEAGI VGAYLAEKEAGI VGAYLVKKEAAI VGAYLVVKEAGI VGAYLVQTEVGI VGAYLVQTEVGI CGAYYVKEEVGI CGAYYVKEEVGI CGAYYVKEEVGI CGAYYVKEEUGI VGAELVKRELGI VGAELVKRELGI VGAEVKELGI VGAPVKELGI VGAEVKELGI VGAEVKELGI VGAEVKELGI VGIYKIQEDLAI VGIYKIQEDLAI VGIYKIQEDLAI VGIYKIQEDLGL VGYKIQEDLGI AAAEIAERELGI	QDEDIIDAIRYHTS TDPEILDAIRYHTS TDPEILDAIRYHTS TDPEILDAIRYHTS TDPEILDAIRYHTS TDPEILDAIRYHTS DDPEVLDAIRYHTS EDEDILAIRYHTS EDEDILSAIRYHTS EDEDILAAIRYHT TDADILQAIRFHTT TDADILQAIRFHTT TDADILQAIRFHTT TDADILQAIRIHTT TDESILQAIRIHTT TDESILQAIRLHTT TDESILQAIRLHTT TDESILQAIRLHTT TDESILQAIRLHTT TDESILQAIRLHTT TDESILQAIRAHTV KDQDILAAIAKHTV KDQDILAAIAKHTV KDQDILAAIAKHTV KDQDILAAIAKHTV KDQEILRSIEIHTV TDSEILRSIEIHTV TDSEILRSIEIHTV	G R P G M T L L E K V I Y V A D S R P G M T L L D K V I Y V A D S R P G M T L L D K V I Y V A D S R P G M T L L D K V I Y V A D S R P G M T L L E K V I Y V A D S R P M T L L E K V I Y V A D S R P M T L L E K V I Y V A D S R N M T L E K V V Y V A D S R N M T L L E K V V Y V A D S R N M T L L E K V V Y V A D S R N M T L L D V V Y V A D S A S M S T L D K I I Y V A D S A S M S T L D K I I Y V A D S A S M S T L D K I Y V A D S A Q M S T L D K I Y V A D S A Q M S T L D K I Y V A D S A Q M S T L D K I Y V A D S A Q M S T L D K I Y V A D S A Q M S T L D K I Y V A D S A Q M S T L D K I Y V A D S A Q M S T L D K I Y V A D S S D M S L D K I Y V A D S H M S T L D K I Y V A D S H M S T L D K I Y V A D S H M S L D K I Y Y A D S K E M S L T A K V Y L A D	А А
B.subtilis	$ \begin{array}{c} \alpha 8 & \alpha 9 \\ 0.0000000 & 0.00000000000 \\ \star 140 & 150 & \star 160 \\ \end{array} $	<u> </u>	α10 0000000000000 180		
B.subtilis B.amyl.M27 B.amyl.M27 B.amyl.TT-45 B.pumilus B.methanolicus G.kaustophilus GeobacillusWCH7 B.licheniformis B.halodurans B.halodurans B.sbiricum E.sibiricum E.sibiricum E.gallinarum T.halophilus M.plutonius CarnobacteriumA L.saerimmeri S.agalactiae S.agalactiae S.salivarius L.buchneri S.termitidis	IAF P GVD EVB KLAETDLIN QA LI QS I KNTT RHFP GVE EVR RLAETDLIN QA LI QAM KNTT RHFP GVE EVR RLAETDLIN EA LI QAM KNTT ROFP GVE EVR RLAE QDLIN EA LI QAM KNTT RHFP GVE EVR RLAE QDLIN EA LI QA KNTT RFP GVD EVR RLAE QDLIN KA LI RS I QNTT RFP GVD EVR RLAE DDLIN KA LI RS I QNTT RFP GVD EVR RLAE DDLIN KA LI KA LQNT RFP GVD EVR RLAE BDLIN KA LI KA LQNT RFP GVE EVR RLAE KDLID EA II SSL VNT ROFP GVE KVR TQA KTDLIN GA II SSL VNT ROFP GVE KVR TA KA LDLIN GA LI SSL VNT ROFP GVE KVR TA KA LDLIN GA VI ST KKNT ROFP GVE KNR TA KA LDLID AA LA LG GT ROFP GVE KAR AT AK KDLID EA VA YE KKNT ROFP GVE EA RLAA KDLID EA VA YE KKNT ROFP GVE EA RLAA KDLID DA VA YE T KNTT ROFP GVE EA RLAAK VDLIN KA VA YE TA RT RFP LVE EA RU LAK VDLIN KA VA YE TA RT RFP LVE EA RU LAK SLD KA VA YE T KNTT RFP PLVE EA RU LAK SLD KA VA YE T KNTT RFP PLVE EA RU LAK SLD KA VA YE T KNTT RFP PLVE EA RU LAK SLD KA VA YE T KNTT RFP PLVE EA RU LAK SLD KA VA YE T KNTT RFP PLVE EA RU LAK SLD KA VA YE T KNTT RFP PLVE EA RU LAK SLD KA VA YE T KNTT RFP PLVE EA RU LAK SLD KA VA YE T KNTT RFP PLVE EA RU TAK SLD KA VA YE T KNTT RFP PLVE EA RU TAK SLD KA VA YE T KNTT RFP PLVE EA RU TAK SLD KA VA YE T KNTT RFP PLVE EA RU TAK SLD KA VA YE T KNTT RFF PLVE EA RU TAK SLD KA VA YE T KNTT RFF PLVE EA RU TAK SLD KA VA YE T KNTT RFF PLVE EA RU TAK SLD KA VA YE T KNTT RFF PLVE EA RU TAK SLD KA VA YE T KNTT RFF PLVE EA RU TAK SLD KA VA YE T KNTT RFF PLVE EA RU TAK SLD KA VA YE T KNTT RFF PLVE EA RU TAK SLD KA VA YE T KNTT RFF PLVE EA RU TAK SLD KA VA YE T KNTT RFF PLVE EA RU TAK SLD KA VA YE T KNTT RFF PLVE KA RU TAK SLD KA VA YE T KNTT RFF PLVE KA RU TAK SLD KA VA YE T KNTT RFF PLVE KA RU TAK SLD KA VA YE T KNTT RFF PLVE KA RU TAK SLD KA VA YE T KNTT RFF PLVE KA RU TAK SLD KA VA YE T KNTT RFF PLVE KA RU TAK SLD KA VA YE T KNTT RFF PLVE KA RU TAK SLD KA VA Y	MV FLMKKNQ PVF MIFLMKKNQ PVF MIFLMKKNQ PVF MIFLMKKNQ PVF IIFLMKKNQ PIY IIFLLEKRQLIY IIFLLEKRQLIY IIFLLEKKQLIY IIFLISKNQAVY IIFFLSKNQAVY IIFFLSKNQPY IIFLISKQVVY IIFFLSKNQPY UFFLEKQTVF LLYLIEKEAPIY LLHLIENERKVY LQFLMEQNEPVY VAFLASKAQPIY VSFLASKAQPIY VSFLASKAQPIY VSFLASKAQPIY VSFLAHQALPIF VEHLAHQALPIF VEHLAHQALPIF IISYLVKASLIH	PDTFLTYNWLVVSGX PETFATYNWLVVSQK PETFATYNWLVVSQK PETFATYNWLVVSQK PDTLATYNSLVHKN PDTLHTYNSLVHKN PDTIHTYNSLVTIK PDTIHTYNSLWNEI PDTLITYNQLLLEQ PDTLITYNQLLLEQ PDTLATYNQLLLEQ PDTLATYNQLLLEQ PDTLATYNALVNDR PLTIKTYNAFVPEK PKTIETYNAWVANK PKSLETYNEWVVKY PKSLETYNEWVVKY PKSLETYNEWVAKK PKSLETYNEWVAKK PKSLETYNEWVAKK PKSLETYNAYIPYL PQTILTYNAVPKP PQTLETYNAFVHYL PQTILTYNSWVPKF PQTILTYNSWVPKF PNTLAFRNDLLTKK	RQEEYGK RQEEYGK N ED KGEAKR. KGECEE. H RGGLNK RGGLNK RGGF. RKGTVL. N KKEPVL. KKEPVL. KKERVKI	

Figure S12. Multiple sequence alignment of the proteins in the YqeK clade. Residues that coordinate the metal ions are indicated with purple arrows.

$\texttt{Thermotog.}_\texttt{ThpO}$	1	αl 2222 223 * 10	α2 200000000 **20	30	200000000 40*	α3 *****	222 50	
Thermotog. Thp0 MesFr Thp0 Kos01 Thp0 ThePe-Thp0 PyrYa Thp0 Thermoc. Thp0 MetKa Thp0 HalLa Thp0 EscCo YedJ ShiFle YedJ CitKo YedJ AzoBr YedJ CitKo YedJ RhoRu YedJ SorCe_YedJ	MNLTE: MDFDR: MLSYSVY: MLLEVVE MLEG.VVKEFLG M. RLEDFIR: MKVNPK. MDLQ. MDLQ. MDLQ. MDLQ. MELQ. MDLLQ. MDLH. MFKDVVLSDVD. MSESSG. MTDDPT. MRREMK.	LMERIPHIK LINCIPHIK LINCIPHIK GEMLR EVN.EIVS EEDSIRLIE FPWBR FPWBR HWQA CHWQA CHWQA CVWES CDWSR LWAQ LWSR	LINIVREAFKI SILGVVEKAFEI SULTEAKRFFE VAEEAVKF ACEMARQS. TREFARGFF. VIEFVRTEMS VRERARYF PENWLKNHHQ PFENWLKNHHQ PFENWLKNHHQ FFENWLKNHHQ SFENWLKNHHQ SITTWLRSNA. RLEEAVVGLGA FAKKLEETDL	DYDDFAHD NYDDFAHD YNGSHHDI LGEGHGI EXPSFHDI DAPPAHD HQDAAHD HQDAAHD HQDAAHD HQDAAHD HQDAAHD HQDAAHD HQDAAHD HQDAAHD HQDAAHD HQDAAHD AEDFAHD ABDGAHD NGDRSHD GKDFAHJ	ISHTFRVNEN SWITERVTES SWITERVTES SHVERVYRI LSHVERVYRI SHTERVLRI FEHVKRVEGI VNHVRRVEGI UCHFRRVWAT ICHFRRVWAT ICHFRRVWGT IFHFRRVWGT IDHLLRVWGT LMLLRVWQT LAHAKRVLRT LAHFARVVAT	A S E T A A S E T A A S E T A A S E T A	SRE	CD JAN EVKH JAD SAD JAD SPDPVAD DVD DAD DAD DAD
	α4	α.5	🖌 α6	η1		α7	¥	
ThermotogThpO	00000000000000 ± 0	0000000 70	8020000 80**	20000000 ★90	··e 2000	20000000 * 11	0 120	ومعدوه
Thermotog. Thp0 MesPr Thp0 Kos01 Thp0 ThePe Thp0 PyrYa_Thp0 Thermoc. Thp0 MetKa Thp0 Halla Thp0 EscCo_YedJ ShiBo_YedJ ShiPle YedJ CitKo YedJ AzoBr YedJ DecSu_YedJ RhoRu_YedJ SorCe_YedJ	$ \begin{array}{l} L \bigcirc KA \\ I \square A \\ L \models I \land V \square A \\ L \vdash D \\ V \square A \\ L \vdash D \\ V \square A \\ A \\ L \vdash D \\ L \vdash V \square A \\ A \\ A \\ I \\ L \\ E \\ V \\ L \\ A \\ A \\ I \\ I$	IKRPHEALTC IKRPEEDFTC IKRKDEVITC VARSMEDEGI IARPLEDAGI IGRPAEERSG VGRAKEDRGI IVSLAKNHPC IVSLAKNHPC IVSLAKNHPC IVSLAKNHPC IVSLAKNHPC IVSLYKNHPI IVSLYKNHPI IVSLYKNHPI IVSLYKNHPI	JV. DHAESGAEK JI. DHAESGAEK JI. DHAESGAEK VEDHAREGAR (IEDHAAEGAR (IEDHAAVEGAK J. EDHAEVSAE SIDDHAEVSAE SIDDHAEVSAE RORSSILAAEJ DRORSSILAAEJ DRYRSSVLAADJ DRSRASTLAAEJ RSSLASRLSAEJ	YASGLUPT YAFQLUAK LAGKLUAK LAGKLUAK LARRETTS SAERLUPE SAERLUPE STRRLURE STRRLURE STRRLURE AARRIURE LAVAK RE KARELUGG	. MGFDISFV . MGFESQFT . HGFDEKTI . LGYPEEKV . LGYPEEKV . LGYPEEKV . LGYPEEKV . LGPDDVV EFEQFPAEKI EFEQFPAEKI DFPAYPEAKI . AGFPAEKI . AGFPAEKI . LDFPAGRC . VEYPEAFL	A EVSKAIR FEVSEAIK EVVAYCIE EKVAYCIE EKVAHAIE EKVAHAIE EAVCHAIAIE EAVCHAIA EAVCHAIA EAVCHAIA AAVSHAIE PATRHAIE DGVAHAIE DGVAHAIE	SHRYSGGLTPT VHRYSGGLNPS SHRYSLKTP AHRYSGGLRP AHRFSGSLKP AHRFSGFEP AHRFSGTGEEP AHRFSTGPEP AHRFSTGPEP AHRYSTDPDP AHSFSAQIAPI AHSFSAQIAPI AHSFSAQIAPI AHSFSAGIAPI AHSFSAGIAPI AHSFSAGIAPI AHSFSAGIAPI	SLTGKI SSMSARI SSMEAQI TLEAKI TTEAKI TTEAKI TTEAKI TTEAKI TTEAKI TTEAKI TIEAKI SLEARI TLEAKI
ThermotogThp0	α8 η2 202020200 0 130 1	α9 <u>00000000</u> 40	L50 10	5 0	170	180		тт
Thermotog. ThpO Thermotog. ThpO MesPr ThpO KosOl ThpO ThePeThpO PyrYa ThPO Thermoc. ThpO MetKa ThpO HalLa ThpO EscCO YedJ ShiBO YedJ ShiPle YedJ CitKo YedJ AzoBr YedJ PcSU YedJ RhoRu YedJ SorCe_YedJ	CAS COLORIZACIÓN CODADRIDAIGA LODADRIDAIGA LODADRIDAIGA LODADRIDAIGA LODADRIDAIGA LODADRIDAIGA LODADRIDAIGA LODADNIDALGA VODADRIEALGA VODADRIEALGA VODADRIESLGA VODADRIDAIGA	09 VAIARVFSY2 IAVARVFSY5 IAVARVFSY5 IAIARVFAH IGVARAFARC IGIARVFMY2 IGIARVFSY0 IGLARVFSY0 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5 IGLARVFAV5	LSO LGVRFYDD GARGTPFYDP GARGTPFYDP GARGTPFYDP GBRRGTPFYDP GBRRGTSLES GENRGRDIEA. GERGTSLES GALGVALFDG GGLGVSLFDA GALGUALFDG GGLGVSLFDA GRLGSALFHG GLLGVRFYDD	50 QFSPRSSYS KLKPGDSY KIKPSKNYI SKPPKERYI DFFAQHRI EDPFAQHRI EDPFAQHRI EDPFAQHRI EDPFAQRRI EDPFAQRRI EDPFARTRI DDPLGTARRI DDPLGTARRI GDPFAERRI	170 SGNSRSI VGK.SSTAI VGK.SSTYI OGH.SETVV OGH.SETVV EAHPEGESQY PLD.DKRYAI PLD.DKRYAI PLD.DKRYAI PLD.DKRYAI SLD.DKQFAI SLD.DKAYAI SLD.DKAYAI SLD.DKAYAI SLD.DRAYAI PLD.ERAYAI PLD.ERAYAI	180 NHFHEXIL NHFFEXIL NHFFEXIL NHFFEXIL RHFEEXIL RHFEEXIL DHFQTXLL DHFQTXLL DHFQTXLL DHFQTXLL DHFQTXLL DHFYTXLL DHFKVXLF	KTRP. KITP. KIDP. KVK. KLR. KLR. KLP. KLP. KLP. KLP. KLP. TLP. TLP. TLP. TLP. TLP. TLP. TLP. T	TT ET ES DT DT DL EL CT QT MQTMQT QT AT AT ES
ThermotogThpO MesPr_ThpO KosOl_ThpO ThePe-ThpO PyrYa_ThpO ThePe-ThpO PyrYa_ThpO HalLa_ThpO HalLa_ThpO EscCo_YedJ ShiBo_YedJ ShiBo_YedJ ShiBo_YedJ CitKo YedJ AzoBr_YedJ RhoRu_YedJ SorCe_YedJ ThermotogThpO	CAS COLORODOLO 130 130 100 100 100 100 100 100	α9 VAIARVFSF2 IAVARVFSF2 IAVARVFSF2 IAIARVFAVG IGIARVFMY2 IGIARVFMY2 IGIARVFMY2 IGLARVFAV2 IGLAVAV2 IGL	SO 1 GKTGTPLHSL GKTGTPLHSP GEKTGTPLYDP GERGTPFYDP GERGTSLES. GEMGSPIHGS. GENGSPIHGS. GLGVALFDG GALGVALFDG GALGVALFDG GLGVSLFDA GLGSALFHD GLLGVRFYDP GUOUDOO	5 Q QFSPRSSYS KLKPGDSYK KLKPSKNYN SKPPKERYI SKPPKERYI EDPFAQHRI EDPFAQHRI EDPFAQHRI EDPFAQHRI EDPFAQRRI EDPFAQRRI EDPFARTRI DDPLAFRRI SDPLAFRRI	170 SGNSRSI (GYSSTAI OGH.SSTYI OGH.SSTYI DGH.SSTYI DGH.SSTYI PLD.DKRYAI PLD.DKRYAI PLD.DKRYAI PLD.DLYAI PLD.DLYAI PLD.DLYAI PLD.CAYAI PLD.CAYAI PLD.CAYAI PLD.CAYAI	180 NHFFBKIL NHFFBKIL NHFFBKIL NHFFBKIL CHIHBKIL GHIHBKIL GHIHBKIL DHFQTKLL DHFQTKLL DHFQTKLL DHFYTKLL DHFYTKLL DHFYTKLL	KTRP. KITP. KITP. KVK. KLR. KLR. KLP. KLP. KLP. KLP. TLP. TLP. TLP. TLP. TLP. TLP. TLP. T	TT ES DT DT

Figure S13. Multiple sequence alignment of the proteins in the ThpO and YedJ clades. Residues that coordinate the metal ions are indicated with purple arrows. Purple shaded boxes indicate the 5th His ligands in ThpO and YedJ respectively.

Table S1. Data collection and refinement statistics for the citrate-bound and substrate-bound(OH-AEP) PhnZ X-ray crystal structures.

	PhnZ-citrate	PhnZ-citrate (ano)	PhnZ-substrate
Data collection			
Wavelength (Å)	0.98	1.72	0.98
Space group	<i>P</i> 3 ₁ 21	<i>P</i> 3 ₁ 21	C2
Unit cell dimensions			
a, b, c (Å)	68.18, 68.18, 98.09	67.97, 67.97, 98.30	148.75, 165.42, 109.35
Resolution (Å)	50.00-1.85 (1.88-1.85)*	50.00-2.92 (2.97-2.92)*	30.00-2.98 (3.05-2.98)*
R _{sym} or R _{merge}	0.063 (0.817)	0.343 (0.163)	0.126 (0.372)
Ι / σ(Ι)	49.6 (3.1)	16.4 (36.2)	13.06 (4.16)
Completeness (%)	99.94 (99.40)	100 (100)	93.5 (96.9)
Redundancy	10.8 (9.8)	20.7 (20.1)	5.9 (5.6)
Refinement			
Resolution range (A)	50.00-1.85		30.00-2.98
No. unique reflections	23752		16580
R _{work} / R _{free}	0.1875/0.2118		0.278/0.333
No. atoms	1593		5767
Protein	1434		5727
Ligand/ions	14		40
Water	154		N/A
Avg. B-factor			
Protein	34.20		67.40
Ligand/ions	26.70		52.80
Water	43.20		N/A
r.m.s deviations			
Bond lengths (Å)	0.006		0.008
Bond angles (°)	0.94		0.452

*Values in parentheses report statistics for the highest-resolution shell

Accession code	NCBI Annotation	Organism
ACU83550.1	predicted HD phosphohydrolase PhnZ	uncultured bacterium HF130_AEPn_1
YP_001091480.1	hypothetical protein P9301_12561	Prochlorococcus marinus str. MIT 9301
YP_001017139.1	hypothetical protein P9303_11251	Prochlorococcus marinus str. MIT 9303
YP_481098.1	hypothetical protein Francci3_1995	Frankia sp. Ccl3
ZP_08717929.1	hypothetical protein MCOL_20456	Mycobacterium colombiense CECT 3035
ZP_01905457.1	HDIG domain protein	Plesiocystis pacifica SIR-1
ZP_01854986.1	hypothetical protein PM8797T_07614	Planctomyces maris DSM 8797
XP_001399948.1	hypothetical protein ANI_1_2760024	Aspergillus niger CBS 513.88
XP_002479453.1	conserved hypothetical protein	Talaromyces stipitatus ATCC 10500
XP_002565218.1	hypothetical protein Pc22g12750	Penicillium chrysogenum Wisconsin 54-1255
YP_006563750.1	metal-dependent phosphohydrolases-like protein	Phaeobacter gallaeciensis 2.10
ZP_05080776.1	conserved hypothetical protein	Rhodobacterales bacterium Y4I
ZP_05786163.1	conserved hypothetical protein	Silicibacter lacuscaerulensis ITI-1157
ZP_05087911.1	conserved hypothetical protein	<i>Ruegeria</i> sp. R11
YP_004014535.1	phosphonate degradation operons associated HDIG domain- containing protein	<i>Frankia</i> sp. Eul1c
YP_004142411.1	hypothetical protein	Mesorhizobium ciceri biovar biserrulae WSM1271
ZP_01074753.1	hypothetical protein MED121_17544	Marinomonas sp. MED121
ZP_01754854.1	hypothetical protein RSK20926_22209	Roseobacter sp. SK209-2-6
ZP_10649542.1	phosphonate degradation operons associated HDIG domain protein	Pseudomonas sp. GM50
ZP_11114572.1	phosphonate degradation operons associated HDIG domain protein	Pseudomonas mandelii JR-1
YP_007032156.1	metal-dependent phosphohydrolase	Pseudomonas sp. UW4
ZP_18348829.1	phosphonate degradation operons associated HDIG domain protein	Pseudomonas fluorescens R124
YP_262944.1	phosphonate degradation operons associated HDIG domain protein	Pseudomonas protegens Pf-5
ZP_10174257.1	phosphonate degradation operons associated HDIG domain protein	Pseudomonas chlororaphis O6
ZP_10142158.1	phosphonate degradation operons associated HDIG domain protein	Pseudomonas synxantha BG33R
AAV65817.1	myo-inositol oxygenase	Rattus norvegicus
XP_003509608.1	PREDICTED: inositol oxygenase-like	Cricetulus griseus
AAF25204.1	myo-inositol oxygenase	Homo sapiens
NP_001124754.1	inositol oxygenase	Pongo abelii
XP_003281515.1	PREDICTED: inositol oxygenase-like	Nomascus leucogenys
XP_003461613.1	PREDICTED: inositol oxygenase-like	Cavia porcellus
XP_001490312.3	PREDICTED: inositol oxygenase-like	Equus caballus(horse)
AAK00766.1	kidney-specific protein 32	Homo sapiens
CAJ75704.1	predicted HD superfamily hydrolase	uncultured Thermotogales bacterium
AFK07215.1	putative domain HDIG-containing protein	Mesotoga prima MesG1.Ag.4.2
YP_002940461.1	metal-dependent phosphohydrolase	Kosmotoga olearia TBF 19.5.1
YP_920563.1	metal-dependent phosphohydrolase	Thermofilum pendens Hrk 5
NP_614753.1	HD superfamily hydrolase	Methanopyrus kandleri AV19
YP_004624251.1	metal-dependent phosphohydrolase, HD superfamily	Pyrococcus yayanosii CH1
YP_006424365.1	hypothetical protein containing HD domain 1	Thermococcus sp. CL1
ZP_09946302.1	metal-dependent phosphohydrolase	Halobiforma lacisalsi AJ5
NP_416471.1	predicted hydrolase, HD superfamily	Escherichia coli str. K-12 substr. MG1655
YP_407527.1	hypothetical protein SBO_1046	Shigella boydii Sb227
EGM61578.1	HD domain protein	Shigella flexneri J1713
YP_001452563.1	hypothetical protein CKO_00980	Citrobacter koseri ATCC BAA-895
CCD01718.1	putative metal-dependent phosphohydrolase	Azospirillum brasilense Sp245
YP_005029269.1	unnamed protein product	Dechlorosoma suillum PS

Table S2. Proteins used for the phylogenetic analysis.

YP_001615272.1 YP 427146.1 P54456.1 YP 148374.1 YP 079904.1 YP_001487526.1 YP_002771473.1 YP 001813266.1 ZP 05648832.1 ZP_03943580.1 NP 242193.1 YP_003427593.1 YP 002950417.1 ZP_11547518.1 ZP 16170044.1 ZP 17179708.1 YP_003921057.1 YP 003307575.1 NP 688652.1 NP 736140.1 YP 005095370.1 ZP_12658042.1 YP_004886756.1 ZP 19196683.1 YP 004455861.1 ZP 02184981.1 NP 968683.1 YP 005036194.1 YP_007644848.1 YP 004196908.1 YP_004119802.1 ZP 07334137.1 NP 864187.1 EGF24277.1 NP_233069.1 YP_005632344.1 YP 002395757.1 WP 005508357.1 YP_868720.1 YP 933792.1 YP_423590.1 YP_001996362.1 NP 416794.1 EIE57001.1 EHS94221.1 YP 003742460.1 YP_005401021.1 AFH92378.1 YP 005634320.1 YP 002506133.1

unnamed protein product metal-dependent phosphohydrolase, HD region Uncharacterized protein yqeK hypothetical protein GK2521 hypothetical protein BL02082 yqeK gene product hypothetical protein BBR47_19920 metal-dependent phosphohydrolase HD domain-containing protein HD superfamily hydrolase hypothetical protein BH1327 hypothetical protein BpOF4_13250 metal-dependent phosphohydrolase YqeK 2',3'-cyclic-nucleotide 2'-phosphodiesterase YqeK hydrolase kidney-specific protein 32 hypothetical protein SAG1661 hypothetical protein gbs1705 hydrolase putative HD superfamily hydrolase hypothetical protein TEH_12650 hypothetical protein D271_06905 HAD superfamily hydrolase HD domain protein hypothetical protein Bd1817 unnamed protein product HD-GYP hydrolase domain-containing protein metal-dependent phosphohydrolase metal-dependent phosphohydrolase metal-dependent phosphohydrolase response regulator HD-GYP domain-containing protein hypothetical protein VCA0681 HDIG domain-containing protein hypothetical protein VS II1179 hypothetical protein metal-dependent phosphohydrolase HD-domain-containing protein HD-GYP domain-containing protein metal dependent phosphohydrolase 5'-nucleotidase YfbR 5'-nucleotidase YfbR UPF0207 protein YfbR unnamed protein product unnamed protein product 5'-nucleotidase unnamed protein product unnamed protein product

Sorangium cellulosum 'So ce 56' Rhodospirillum rubrum ATCC 11170 Bacillus subtilis subsp. subtilis str. 168 Geobacillus kaustophilus HTA426 Bacillus licheniformis ATCC 14580 Bacillus pumilus SAFR-032 Brevibacillus brevis NBRC 100599 Exiguobacterium sibiricum 255-15 Enterococcus gallinarum EG2 Lactobacillus buchneri ATCC 11577 Bacillus halodurans C-125 Bacillus pseudofirmus OF4 Geobacillus sp. WCH70 Bacillus methanolicus MGA3 Bacillus amyloliquefaciens subsp. plantarum M27 Bacillus amyloliquefaciens IT-45 Bacillus amyloliquefaciens DSM 7 Sebaldella termitidis ATCC 33386 Streptococcus agalactiae 2603V/R Streptococcus agalactiae NEM316 Streptococcus macedonicus ACA-DC 198 Streptococcus salivarius M18 Tetragenococcus halophilus NBRC 12172 Lactobacillus saerimneri 30a Melissococcus plutonius ATCC 35311 Carnobacterium sp. AT7 Bdellovibrio bacteriovorus HD100 Bacteriovorax marinus SJ Bdellovibrio exovorus JSS Geobacter sp. M18 Desulfovibrio aespoeensis Aspo-2 Desulfovibrio fructosovorans JJ Rhodopirellula baltica SH 1 Rhodopirellula baltica WH47 Vibrio cholerae O1 biovar El Tor str. N16961 Vibrio cholerae LMA3984-4 Vibrio splendidus LGP32 Vibrio mimicus Shewanella sp. ANA-3 Azoarcus sp. BH72 Magnetospirillum magneticum AMB-1 Chloroherpeton thalassium ATCC 35110 Escherichia coli str. K-12 substr. MG1655 Escherichia coli Al27 Klebsiella oxytoca 10-5245 Erwinia billingiae Eb661 Rahnella aquatilis HX2 Providencia stuartii MRSN 2154 Vibrio cholerae LMA3984-4 Clostridium cellulolyticum H10

REFERENCES

- 1. Gasteiger E. HC, Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook*, ed Walker JM (Humana Press), pp 571-607.
- 2. Bollinger JM, Jr. (1993) On the chemical mechanism of assembly of the tyrosyl radicaldinuclear iron cluster cofactor of *E.coli* ribonucleotide reductase. Ph. D. (Massachusetts Institute of Technology).
- 3. Dassama LMK, Boal AK, Krebs C, Rosenzweig AC, & Bollinger JM, Jr. (2012) Evidence that the β subunit of *Chlamydia trachomatis* ribonucleotide reductase is active with the manganese ion of its manganese(IV)/iron(III) cofactor in site 1. *J. Am. Chem. Soc.* 134(5):2520-2523.
- 4. Hammerschmidt F & Völlenkle H (1989) Absolute Konfiguration der (2-Amino-1hydroxyethyl)phosphonsäure aus *Acanthamoeba castellanii* (Neff) - Darstellung der Phosphonsäure-Analoga von (+)- und (-)-Serin. *Liebigs Ann. Chem.* 1989(6):577-583.
- 5. Bollinger JM, Jr., *et al.* (1995) Use of rapid kinetics methods to study the assembly of the diferric-tyrosyl radical cofactor of *E. coli* ribonucleotide reductase. *Methods Enzymol.* 258:278-303.
- 6. Münck E (2000) Aspects of ⁵⁷Fe Mössbauer spectroscopy. *Physical Methods in Bioinorganic Chemistry*, ed Que L, Jr. (University Science Books, Sausalito, CA), pp 287-319.
- 7. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. 276:307-326.
- 8. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66(Pt 4):486-501.
- 9. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66(Pt 2):213-221.
- 10. Vonrhein C, Blanc E, Roversi P, & Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol. Biol.* 364:215-230.
- 11. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* 62(Pt 9):1002-1011.
- 12. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, & Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr.* D61(4):458-464.
- 13. Schüttelkopf AW & van Aalten DM (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* 60(Pt 8):1355-1363.
- 14. Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ (1990) Basic local alignment search tool. *J. Mol. Biol.* 215(3):403-410.
- 15. Katoh K & Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30(4):772-780.
- 16. Thompson JD, Higgins DG, & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22(22):4673-4680.

- 17. Tamura K, *et al.* (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28(10):2731-2739.
- 18. Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4(4):406-425.
- 19. Takahashi K & Nei M (2000) Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution, and maximum likelihood when a large number of sequences are used. *Mol. Biol. Evol.* 17(8):1251-1258.
- 20. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947-2948.