



NIH Public Access

Author Manuscript

Biochemistry. Author manuscript; available in PMC 2013 February 28.

Published in final edited form as:
Biochemistry. 2012 February 28; 51(8): 1762–1773. doi:10.1021/bi201838b.

Structure-Based Function Discovery of an Enzyme for the Hydrolysis of Phosphorylated Sugar Lactones

Dao Feng Xiang[‡], Peter Kolb[§], Alexander A. Fedorov^Ψ, Chengfu Xu[‡], Elena V. Fedorov^Ψ, Tamari Narindoshvili[‡], Howard J. Williams[‡], Brian K. Shoichet^{§,*}, Steven C. Almo^{Ψ,*}, and Frank M. Rauschel^{‡,*}

[‡]Department of Chemistry, P.O. Box 30012, Texas A&M University, College Station, Texas 77842-3012

[§]Department of Pharmaceutical Chemistry, University of California, San Francisco, 1700 4th Street, San Francisco, California 94158-2330

^ΨAlbert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

Abstract

Two enzymes of unknown function from the cog1735 subset of the amidohydrolase superfamily (AHS), LMOf2365_2620 (Lmo2620) from *Listeria monocytogenes* str. 4b F2365 and Bh0225 from *Bacillus halodurans* C-125, were cloned, expressed and purified to homogeneity. The catalytic functions of these two enzymes were interrogated by an integrated strategy encompassing bioinformatics, computational docking to three-dimensional crystal structures, and library screening. The three-dimensional structure of Lmo2620 was determined at a resolution of 1.6 Å with two phosphates and a binuclear zinc center in the active site. The proximal phosphate bridges the binuclear metal center and is 7.1 Å away from the distal phosphate. The distal phosphate hydrogen bonds with Lys-242, Lys-244, Arg-275 and Tyr-278. Enzymes within cog1735 of the AHS have previously been shown to catalyze the hydrolysis of substituted lactones.

Computational docking of the high energy intermediate (HEI) form of the KEGG database to the three-dimensional structure of Lmo2620 highly enriched anionic lactones versus other candidate substrates. The active site structure and the computational docking results suggested that probable substrates would likely include phosphorylated sugar lactones. A small library of diacid sugar lactones and phosphorylated sugar lactones was synthesized and tested for substrate activity with Lmo2620 and Bh0225. Two substrates were identified for these enzymes, D-lyxono-1,4-lactone-5-phosphate and L-ribono-1,4-lactone-5-phosphate. The k_{cat}/K_m values for the cobalt-substituted enzymes with these substrates are $\sim 10^5 \text{ M}^{-1} \text{ s}^{-1}$.

*To whom correspondence may be addressed: (FMR) telephone: (979) 845-3373; fax: (979)-845-9452; rauschel@tamu.edu, (SCA) telephone: (718) 430-2746; fax: (718)-430-8565; almo@ecom.yu.edu, (BKS) telephone: (415)-514-4126; fax: (415)-514-4260; shoichet@cgl.ucsf.edu.

ASSOCIATED CONTENT

Supporting Information

A figure showing the enrichment of the most frequently observed reaction types from the docking of the HEI KEGG database to the structure of Lmo2620 (Figure S1), a list of additional lactones tried a potential substrates (Table S1), and a list of proteins expected to have the same substrate profile as Lmo2620 and Bh0225 (Table S2) is available. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Author Contributions

DFX, PK, AAF, and CX contributed equally.

[†]The X-ray coordinates and structure factors for Lmo2620 have been deposited in the Protein Data Bank (PDB accession code: 3PNZ).

Function determination of uncharacterized enzymes remains challenging for the millions of new proteins identified from large scale DNA sequencing efforts since there are currently no reliable high-throughput methods for deciphering enzyme specificity. Thus, the mapping of new reactions and metabolic pathways has not kept pace with identification of new genes within completely sequenced bacterial genomes. An assessment of the functional annotations for the more than 12 million genes thus far sequenced suggests that approximately one-third of the encoded proteins have an uncertain, unknown, or incorrect function (1–5). This implies that a substantial portion of the metabolic landscape remains to be identified. We have therefore explored developing comprehensive strategies for annotating enzymes of unknown function through a combination of bioinformatics, computational docking of potential ligands to protein structures and synthesis of focused compound libraries for experimental verification. We have applied these methods toward the elucidation of function for enzymes within the amidohydrolase superfamily (6–20).

The amidohydrolase superfamily (AHS) was first recognized by Sander and Holm based on the structural similarities of urease, phosphotriesterase, and adenosine deaminase (21). Enzymes within this superfamily catalyze hydrolytic reactions of amide and ester bonds contained within carbohydrates, peptides, and nucleic acids (22). Less common amidohydrolase-catalyzed reactions include decarboxylations, hydrations, and isomerizations (11, 23–25). All proteins within the AHS have a distorted (β/α)₈-barrel fold and contain either a mononuclear or binuclear metal center in the active site. The metal center activates a hydrolytic water molecule for nucleophilic attack and enhances the electrophilic character of the target substrate. Over 12,000 unique protein sequences have been identified as members of the AHS within the genomes of the first 1,000 completely sequenced bacteria deposited in NCBI. Proteins identified in the AHS have been subdivided into 24 Clusters of Orthologous Groups (COG) (26).

Enzymes in cog1735 have been reported to catalyze the hydrolysis of organophosphates and lactones. Shown in Figure 1 is a sequence similarity network analysis representation of this COG at a BLAST E-value cutoff of 1e-80 (27). The bacterial phosphotriesterase (PTE) from group 8 catalyzes the hydrolysis of a variety of organophosphates including the insecticide paraoxon and the chemical warfare agents GB, GD, and VX (28). Sso2522, from the hyperthermophilic *Sulfolobus solfataricus* of group 9, has a weak phosphotriesterase activity but hydrolyzes lactones at substantially faster rates (29). Three other proteins from group 3, Mt0240 from *Mycobacterium tuberculosis* (30), Rer55000 from *Rhodococcus erythropolis* (31), and MAP3668c from *Mycobacterium avium* (32) have been shown to catalyze the hydrolysis of *N*-acyl homoserine lactones. A thermostable lactonase from *Geobacillus stearothermophilus* (gi|258588268) from group 7 has phosphotriesterase and lactonase activities (33). The PTE homology protein (PHP) from *E.coli* K12 (b3379) from group 1 was structurally characterized in 1998 but its catalytic function remains unknown (34). We, and others, have determined the three-dimensional structure and substrate profile of Dr0930 from *Deinococcus radiodurans* (PDB code: 3FDK and 3HTW) from group 7 (17, 35). This protein has very low phosphotriesterase activity but efficiently hydrolyzes δ - and γ -lactones with an alkyl substitution at the carbon adjacent to the ring oxygen.

Here we combine structure-based docking screens of a general metabolite library with biochemical screens of a focused chemical library to determine the substrate profile for enzymes from group 5 of cog1735. LMOf2365_2620 (Lmo2620) from *Listeria monocytogenes* str. 4b F2365 and Bh0225 from *Bacillus halodurans* C-125 were purified to homogeneity and the three-dimensional structure of Lmo2620 determined at a resolution of 1.6 Å. These two proteins share 74% sequence identity and were found to be lactonases that catalyze the hydrolysis of D -lyxono-1,4-lactone-5-phosphate and L -ribono-1,4-lactone-5-phosphate.

MATERIALS and METHODS

Materials

The genomic DNA of *Listeria monocytogenes* str. 4b F2365 and *Bacillus halodurans* C-125 were purchased from ATCC. Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs. The expression vector pET30a (+) and *E.coli* strain BL21 (DE3) were obtained from Novagen and the Platinum *Pfx* DNA polymerase was obtained from Invitrogen. The Wizard Miniprep DNA purification kit was obtained from Promega. LB broth was purchased from Tpi Research Products International. Chromatographic gel filtration columns and Resource Q anion exchange columns were purchased from GE Healthcare. ICP standards were purchased from Inorganic Ventures. All other buffers, purification reagents, and chemicals used in this investigation were purchased from Sigma, unless otherwise stated.

Synthesis of Sugar Lactone Phosphates

The phosphorylated sugar lactones (compounds **1–10**) were made by direct phosphorylation of the parent lactone or the 2,3-isopropylidene protected lactone. D -Ribono-1,4-lactone was converted to 2,3-*O*-isopropylidene- D -ribono-1,4-lactone (36) and then 5.0 mmol of this material was reacted with diphenylchlorophosphate (6.0 mmol) in pyridine to obtain 5-(diphenylphosphate)-2,3-*O*-isopropylidene- D -ribono-1,4-lactone with a yield of 80%. This intermediate (4.0 mmol) was dissolved in 40 mL of acetic acid containing 50 mg of PtO₂ and then H₂ was applied overnight. After removal of the catalyst and solvent, a quantitative yield of 2,3-*O*-isopropylidene- D -ribonolactone-5-phosphate was obtained. The protecting group was removed by the addition of trifluoroacetic acid/dichloromethane (9/1) for 4 hours to obtain D -ribono-1,4-lactone-5-phosphate (**4**). Two equivalents of sodium bicarbonate were added to obtain the disodium salt. ³¹P NMR (121.4 MHz, CDCl₃): 3.36 ppm (5%), 2.54 ppm (95%); MS (ESI negative mode): found: 226.99 (M-H)⁻ and C₅H₉O₈P (M) computed: 228.00. D -Lyxono-1,4-lactone-5-phosphate (**1**) was prepared in the same way from D -lyxono-1,4-lactone (37).

Compounds **2**, **3**, and **5–9** were synthesized similarly but the isopropylidene protection step was not used. For the synthesis of compound **2**, D -xylose (6.0 g, 40 mmol) and NaHCO₃ (5.04 g, 60 mmol) were stirred together in 50 mL water at 0 °C and then Br₂ (3×0.76 mL, 44 mmol) was added at 15–20 minute intervals. The mixture was stirred for 1 hour and then for 3 hours at room temperature. The solvent was removed and the semi-solid residue was extracted with ethanol (2×100 mL). After removal of the ethanol, the oily crude product (D -xylono-1,4-lactone) was obtained in a yield of ~50%. The crude D -xylono-1,4-lactone (~5 mmol) was mixed with diphenylchlorophosphate (6.0 mmol) in 60 mL of pyridine and stirred overnight. The solvent was removed and the product purified by elution from a silica gel column with a mixture of ethyl acetate/acetone (6/1) to obtain 5-diphenylphosphate- D -xylono-1,4-lactone in a yield of 50%. This compound (2.0 mmol) was hydrogenated in acetic acid with a PtO₂ catalyst overnight to obtain D -xylono-1,4-lactone-5-phosphate. The disodium salt was made as described above. ³¹P NMR (121.4 MHz, CDCl₃): 2.96 ppm; MS (ESI negative mode): found: 226.98 (M-H)⁻ and C₅H₉O₈P (M) computed: 228.00. Compounds **3** and **5–9** were obtained using the same procedure with similar results. All of the preparations had a 1–20% contamination of the hydrolyzed lactone. L -Ascorbate-6-phosphate (**10**) was prepared as described previously (38–39). The structures of these compounds are presented in Scheme 1.

Synthesis of Dicarboxylate Sugar Lactones

Compounds **11–18** were synthesized from xylitol, D -arabitol, L -arabitol, and adonitol. The alcohols were oxidized by nitric acid to the corresponding sugar acids and dehydrated to a

mixture of 1,4- and 2,5-lactones. The major contaminant, oxalic acid, was removed at low temperature and then the monoacid lactone converted to the sodium salt with NaHCO_3 . Oxidation of xylitol and conversion of the xylaric acid product to the monolactone provided a racemic mixture of compounds **11** and **15**. A 2:1 mixture of compounds **12** and **16** was prepared by the oxidation of D -arabitol. Oxidation of L -arabitol gave a 2:1 mixture of compounds **13** and **17**. The racemic mixture of compounds **14** and **18** was initiated by the oxidation of adonitol. All four preparations were contaminated with 10–15% of the diacid intermediate. The structures of compounds **11–18** are presented in Scheme 2.

The sugar alcohol (3 g of xylitol, D -arabitol, L -arabitol or adonitol) was dissolved in concentrated nitric acid (12 mL) and then 80 mg of NaNO_2 was added. The mixture was heated in an oil bath at 50 °C for 5–10 minutes. After the evolution of gas subsided, the solution was heated at 50–60 °C for an additional 16 hours. The reaction mixture was allowed to cool to room temperature and then further cooled to –20 °C. The white crystalline precipitate (oxalic acid) was filtered and the clear solution was cooled and filtered again. The solvent was removed by evaporation under reduced pressure at 70–75 °C for 5–6 hours. This procedure allows for the removal of nitric acid and lactonization of acid sugars. The resulting products were dried for 24 hours under vacuum at room temperature as colorless or light yellow solids. The lactone products were obtained in the yields of 50–60%.

Xylaric Acid 1,4- and 2,5-Lactones (**11**, **15**)

^1H NMR (D_2O): δ 4.41 (d, $J = 8.1$ Hz, 1H), 4.55 (t, $J = 7.7$ Hz, 1H), 5.01(d, $J = 7.5$ Hz, 1H). ^{13}C NMR (D_2O): δ 70.96, 72.55, 77.21, 171.01, 175.91. TOF MS/ESI (M-H) $^-$ 161.47 (found), 162.02 (M, calculated).

D-Arabinaric Acid 1,4- and 2,5-Lactones (**12**, **16**)

^1H NMR (D_2O): δ 4.26 (t, $J = 8.5$ Hz, 0.5 H), 4.43 (d, $J = 8.7$ Hz, 0.5H), 4.55–4.65 (m, 2,5 H), 5.00 (d, $J = 3.0$ Hz). TOF MS/ESI (M-H) $^-$ 161.02 (found), 162.02 (M, calculated). D-arabinaric acid (1,4)-lactone (**12**): ^{13}C NMR (D_2O): δ 69.76, 69.80, 78.24, 169.65, 176.6. D-arabinaric acid (2,5)-lactone (**16**): ^{13}C NMR (D_2O): δ 73.04, 75.20, 76.79, 171.17, 174.5.

L-Arabinaric Acid (1,4)- and 2,5-Lactones (**13**, **17**)

^1H NMR ($\text{DMSO}-d_6$): δ 4.07 (t, $J = 7.5$ Hz, 0.5 H), 4.19 (d, $J = 7.2$ Hz, 0.5H), 4.41(t, $J = 3.3$ Hz, 1H), 4.45 (d, $J = 4.8$ Hz, 1H), 4.93 (d, $J = 3.0$ Hz, 1H). TOF MS/ESI (M-H) $^-$ 161.02 (found), 162.02 (M, calculated).

Ribaric Acid-1,4- and 2,5-Lactones (**14**, **18**)

^1H NMR ($\text{DMSO}-d_6$): δ 4.25 (d, $J = 5.1$ Hz, 1H), 4.31 (d, $J = 4.8$ Hz, 1H), 4.74 (S, 1H). ^{13}C NMR ($\text{DMSO}-d_6$): δ 68.43, 71.24, 81.10, 170.00, 176.11. TOF MS/ESI (M-H) $^-$ 161.08 (found), 162.02 (M, calculated).

A racemic mixture of compounds **19** and **20** was synthesized from mucic acid (40). D-glucaro-1,4-lactone (**21**) and D-glucaro-6,3-lactone (**22**) were synthesized from D-glucaric acid (41). Lactones were converted to their sodium salts with NaHCO_3 .

Cloning of Lmo2620 and Bh0225

The gene for Lmo2620 (gi|46908819) was amplified from the genomic DNA of *Listeria monocytogenes str. 4b F2365*. The amplified PCR product was purified using a PCR-cleanup system (Promega), digested with *AseI* and *BamHI*, and ligated in a pET30a (+) vector that had been digested with *AseI* and *BamHI*. The entire cloned fragment was sequenced to confirm the fidelity of the PCR amplification. Bh0225 (gi|15612788) was

amplified from the genomic DNA of *Bacillus halodurans* C-125. The amplified PCR product was purified using a PCR cleanup system (Promega), digested with AseI and BamHI, and ligated to a pET30a(+) vector that had been digested with AseI and BamHI. The cloned fragment was sequenced to verify the fidelity of the PCR amplification.

Expression and Purification of Lmo2620 and Bh0225

To express Lmo2620, BL21(DE3) cells (Novagen) were transformed with the pET30a(+) plasmid containing the gene for Lmo2620. A single freshly transformed colony was cultured in LB medium supplemented with 50 µg/mL kanamycin at 37 °C. The overnight culture was inoculated into 1 L LB media and cultured at 30 °C with vigorous shaking. When the OD₆₀₀ of the culture reached 0.6, the expression of the target protein was induced with 0.5 mM isopropyl-D-thiogalactopyranoside (IPTG) and 1.0 mM Zn(OAc)₂ was added to the culture. Cells were allowed to grow at room temperature for 18 hours and then harvested by centrifugation (6000 rpm, 4 °C, 15 minutes). For Lmo2620, 10 grams of fresh or frozen cells were resuspended in 50 mL of a solution containing 50 mM Hepes, pH 7.5, and then lysed by sonication (5 second pulses for 30 minutes) at 0 °C. After centrifugation, the nucleic acids were removed by adding 20 mL of a 2% (w/v) solution of protamine sulfate in 50 mM Hepes, pH 7.5. After centrifugation the supernatant solution was fractionated with 40–60% saturation of ammonium sulfate. The precipitated protein was resuspended in 50 mM Hepes, pH 7.5, filtered with a 0.2 µm syringe filter (VWR) and loaded onto a Hiload 26/60 Superdex-200 gel filtration column (Amersham Pharmacia). Protein was eluted at a flow rate of 2.0 mL/minute. The fractions containing the target protein were pooled, loaded onto a Resource Q anion exchange column (Amersham Pharmacia), and then eluted with a linear gradient of NaCl in 20 mM Hepes, pH 7.5. The purity of the protein was estimated to be greater than 95% based on SDS-PAGE. The purified protein was designated as [Zn/Zn]-Lmo2620. Lmo2620 containing manganese or cobalt in the active site was obtained in good yield by adding 1.0 mM MnCl₂ or CoCl₂ during expression. The iron-specific chelator, 2, 2'-bipyridyl, was added at a concentration of 0.1 mM to the growth medium when the absorbance reached 0.1 when the cells were grown in the presence of MnCl₂ to reduce the iron content in the protein (42). These proteins are designated as [Mn/Mn]-Lmo2620 and [Co/Co]-Lmo2620, respectively.

Bh0225 was expressed and purified using the same protocol for Lmo2620. Three forms of Bh0225 were prepared ([Zn/Zn]-Bh0225, [Mn/Mn]-Bh0225 and [Co/Co]-Bh0225) by supplementing the growth medium with 1.0 mM ZnCl₂ MnCl₂ or CoCl₂ at the time of induction. Each form of Bh0225 was purified to homogeneity using ammonium sulfate fractionation (40–60% saturation), gel filtration chromatography, and anion exchange chromatography. The purity of Bh0225 was estimated to be greater than 95% by SDS-PAGE.

Metal Analysis

The protein concentration was determined spectrophotometrically at 280 nm using a SPECTRAmax-340 UV-vis spectrophotometer. The extinction coefficient at 280 nm used for calculating the concentration of Lmo2620 was 47,455 M⁻¹ cm⁻¹ and the extinction coefficient at 280 nm for calculating the concentration of Bh0225 was 46,490 M⁻¹ cm⁻¹ (ca.expasy.org/tools/protparam.html). The metal content of the as-purified proteins was determined by inductively coupled plasma emission mass spectrometry (ICP-MS). Before ICP-MS measurements, the protein samples were treated with concentrated nitric acid for 20 minutes at 100 °C and then diluted with distilled water until the final concentration of nitric acid was 1%.

Kinetic Measurements and Data Analysis

The kinetic measurements were conducted with a SPECTRAmax-340 plate reader. The hydrolysis of lactones was monitored using a pH-sensitive colorimetric assay (17). Protons released from the carboxylate product were measured using the pH-indicator, cresol purple. The reactions were performed in 2.5 mM Bicine, pH 8.3, containing 0.2 M NaCl, variable amounts of substrate, 0.1 mM cresol purple and Lmo2620 or Bh0225 in 1% DMSO. The change in absorbance at 577 nm was monitored. The conversion factor ($\Delta\text{OD}/\text{mol of H}^+$) was determined to be $\epsilon = 1.76 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$ using acetic acid. The protein was stored in 50 mM Hepes, pH 7.5, and the buffer was exchanged to 10 mM Bicine, pH 8.3, using a PD-10 desalting column (GE Healthcare) before use. A background rate was observed in the absence of substrate due to acidification by atmospheric CO₂. The background rate was independent of substrate concentration and subtracted from the initial rates.

Kinetic parameters were obtained by fitting the initial rates directly to equation 1, where v is the initial velocity, E_t is the enzyme concentration, k_{cat} is the turnover number, $[A]$ is the substrate concentration, and K_m is the Michaelis constant.

$$v/E_t = k_{\text{cat}}[A]/(K_m + [A]) \quad (1)$$

Crystallization and Data Collection

Crystals of Lmo2620 complexed with zinc and phosphate were grown by the hanging drop vapor diffusion method at room temperature by mixing equal volumes of protein and precipitant, and equilibrating over the precipitant. The protein solution contained Lmo2620 (13.5 mg/mL) in 20 mM Hepes (pH 7.5), 0.1 M NaCl, and 1.0 mM ZnCl₂; the precipitant contained 0.4 M Na₃PO₄, 1.6 M K₃PO₄, 0.1 M imidazole (pH 8.0), 0.2 M NaCl, and 1.0 mM ZnCl₂. Crystals appeared in 3–4 days and exhibited diffraction consistent with space group P2₁, with six molecules of the protein complex per asymmetric unit (Table 1). Prior to data collection, crystals were transferred to cryoprotectant solutions composed of their mother liquids and 20% glycerol. After incubation for ~10 seconds, the crystals were flash-cooled in a nitrogen stream. Diffraction data were collected at the NSLS X4A beamline (Brookhaven National Laboratory) on an ADSC CCD detector. Diffraction data were integrated and scaled with programs DENZO and SCALEPACK (43). The data collection statistics are given in Table 1.

Structure Determination and Model Refinement

The structure of the Lmo2620·Zn²⁺·PO₄³⁻ crystalline complex was determined by molecular replacement with BALBES (44), using the phosphotriesterase homology protein from *E. coli* (PDB ID: 1BF6) as the search model. The initial model was subjected to several iterative cycles of refinement and model building with COOT (45), PHENIX (46) and ARP (47). The model of the Lmo2620-Zn²⁺·PO₄³⁻ complex was refined at 1.6 Å with an R_{cryst} of 0.178 and an R_{free} of 0.197. The final model contains six protein molecules organized in two trimers. Each protein molecule contains, two well-defined zinc ions, and two bound phosphates. The final crystallographic refinement statistics for the Lmo2620-Zn²⁺·PO₄³⁻ complex are provided in Table 1.

Computational Docking to Lmo2620

Concurrent with the *in vitro* experiments, a computational strategy was used to predict substrates for Lmo2620. The high-energy intermediate (HEI) version of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database of biogenic compounds was docked to the active site (48, 49). A HEI is a transiently stable structure that resembles the transition-state of a substrate in both structure and charge distribution. HEI states were

calculated as described (14), resulting in approximately 22,500 different high-energy forms for the 4,207 molecules in KEGG that are candidates for one of the reactions known to be catalyzed by a member of the amidohydrolase superfamily. Because KEGG contains few phosphorylated lactones, a dedicated library was generated in a second step. This library consisted of all phosphorylated 1,4- and 1,5-lactones originating from pentose and hexose sugars, and the HEI forms were generated as before.

Docking was carried out with DOCK3.5.54 as described (8, 17). The computed poses were subjected to a distance cutoff to make sure that the O⁻ of the HEI portion of the molecule is found within 4 Å of the metal ions in the binding site. Of the molecules that satisfied this constraint, the top 500 compounds ranked by DOCK score (consisting of van der Waals, Poisson-Boltzmann electrostatic, and ligand desolvation penalty terms) were inspected visually to ensure compatibility of the pose with the amidohydrolase reaction mechanism.

RESULTS

Cloning, Expression and Purification of Lmo2620 and Bh0225

The genes for Lmo2620 and Bh0225 were successfully cloned. Lmo2620 expressed well in *E. coli* in the presence of added divalent metal ions and the protein was purified to homogeneity. The metal content of the purified proteins was measured by ICP-MS and the results are shown in Table 2. Each enzyme contained 2.0 ± 0.1 equivalents of metal and in each case the metal added to the growth medium dominated. Minor amounts of iron and/or zinc were also found. The identity of Lmo2620 was confirmed by sequencing the first five amino acids from the N-terminus of purified [Zn/Zn]-Lmo2620. The amino acid sequence (SFIRT) matched that of MSFIRT predicted from the DNA sequence. Bh0225 was expressed in good yield in *E. coli* in the presence of added divalent metal ions. The metal content of the protein samples were measured by ICP-MS and the results are presented in Table 2.

Three-Dimensional Structure of Lmo2620

The three-dimensional structure of [Zn/Zn]-Lmo2620 was determined to a resolution of 1.6 Å. Lmo2620 adopts a distorted (β/α)₇ β -barrel fold with N- and C-terminal extensions (Figure 2). The following chain segments are included in the eight β -strands of the barrel: β -1 (residues 18–28), β -2 (61–64), β -3 (87–94), β -4 (149–156), β -5 (183–187), β -6 (210–214), β -7 (234–237), β -8 (267–270). The N-terminal extension includes a β -loop (3–12) and the C-terminal extension includes two distorted α -helices (290–305, and 311–325). The last helix closes the entrance to the barrel from the N-terminal side. The segment connecting strands β -1 and β -2 of the barrel contains two helices and the long chain segment between strands β -3 and β -4 contains four helices. The remaining segments between the β -strands of the barrel each contain a single helix.

The active site of Lmo2620 is located at the C-terminal end of the barrel and is open to bulk solvent. Two zinc ions are bound in the active site, with the most deeply buried zinc ion and the most solvent exposed zinc ion denoted as α and β , respectively. Zn α is coordinated by His-22 and His-24 from strand β -1, Asp-272 from strand β -8, and the carboxylated Lys-154 from strand β -4 as illustrated in Figure 3. Zn β is coordinated by His-187 from strand β -5 and His-215 from strand β -6, and the carboxylated Lys-154. The two zinc ions are bridged by the carboxylated Lys-154 and a phosphate (proximal) that acts as an additional bridging ligand. The two zinc ions are separated by 3.8 Å, with distances to O3 of the bridging phosphate of 2.0 Å and 2.6–2.8 Å for Zn α and Zn β , respectively. The Zn β is also coordinated by O2 of the bridging phosphate at a distance of 2.3–2.5 Å. Another phosphate (distal) is present at a position near the active site with a P-P distance of 7.1 Å to the proximal

phosphate. There is a water molecule situated between the two phosphates at a distance of 2.5 Å to an oxygen atom of the proximal phosphate and 3.9 Å to an oxygen of the distal phosphate. The shortest distance between the oxygen atoms of the two phosphates is 4.9 Å. The distal phosphate participates in hydrogen bonds with Lys-242, Lys-244, Arg-275, and Tyr-278, as well as an interaction with Arg-218 that is mediated by two water molecules. The distance between the guanidinium amino group of Arg-218 and the distal phosphate is 4.8 Å. The presence of the distal phosphate suggests that Lys-242, Lys-244, Arg-275, Tyr-278 and Arg-218 form a very basic cavity which recognizes substrates possessing a phosphate or carboxylate group.

Docking the HEI Metabolite Library

The KEGG library contains 4,207 metabolites with functional groups recognized by enzymes of the amidohydrolase superfamily. These compounds were docked against the three-dimensional structure of Lmo2620 in an average of 177,707 conformations and 53,990 configurations, with a total calculation time of 362 CPU days. Of the top 500 ranking poses, just under 300 were compatible with the recognized reaction mechanism and five chemotypes were prominent. In order of docking rank these were: lactam hydrolysis, acyclic deamination, acyclic ester and amide hydrolysis, lactone hydrolysis, and exocyclic deamination (Table 3). Detailed examination revealed incomplete fits for most of these compounds. For instance, most of the putative substrates undergoing deamination, amide and ester hydrolysis extended past the active site of Lmo2620, stranding a substantial part of the molecule in bulk solvent. More attractive were the lactones among the top hits, most of which (34/46 hits) had every one of their polar atoms complemented by a corresponding amino-acid interaction with the enzyme. Considering only those molecules that were catalytically competent and that did not strand functionality, these lactones constituted 34/59 of the well-placed metabolites (57.6%, a 10-fold enrichment over the 5.4% of lactones in KEGG) (Table 3 and Figure 5). Even without filtering they made up 46 of the top 500 dock-ranked metabolites. For comparison, there are only four lactam HEIs among the well-positioned poses (out of 815 in the original library). The favorable fits and enrichment of the lactones are also borne out by a ranking that is weighted by the heavy atom count (HAC) of the small molecules—the lactones were typically smaller than many of the other molecules recognized, but more ligand-efficient, and so were among the very top-ranked molecules when the docking score was normalized for molecular size (Supporting Information, Figure S1). Intriguingly, 14 of the lactones were anionic, typically positioned with a carboxylate interacting with Arg-275 and other residues in the structure of Lmo2620 arranged to interact with the distal phosphate (Figure 5B). This observation turned out to be consistent with *in vitro* experiments that were proceeding concurrently, strengthening our confidence in the importance of the lactone chemotype and the importance of a negative charge in this region of the active site.

Substrate Specificity and Kinetic Measurements

Lmo2620 was initially tested for the ability to catalyze the hydrolysis of organophosphate esters. No catalytic activity was detected with paraoxon (diethyl 4-nitrophenyl phosphate), ethyl 4-nitrophenyl phosphate, *bis*(4-nitrophenyl) phosphate, or 4-nitrophenyl phosphate. These results demonstrate that this enzyme is not a generic phosphotriesterase as currently annotated by NCBI. Additionally, Lmo2620 was screened with an array of lactones (Table S1) that have previously been shown to be substrates for enzymes from groups **3** (30–32), **7** (33) and **9** (29) of cog1735 (Figure 1), but no activity was detected with any of these compounds. The most interesting hint for potential substrates of Lmo2620 was the binding of two phosphates in the active site (Figure 3). The phosphate proximal to the metal center is most likely a surrogate for the tetrahedral intermediate formed during substrate hydrolysis, whereas the phosphate distal to the metal center likely mimics a portion of a phosphorylated

(or carboxylated) substrate. Combining this clue with the preponderance of lactone substrates for related proteins from cog1735, suggested the likelihood of phosphorylated lactones as potential substrates. The high-ranking docking compounds did not include lactone phosphates as these are not in the KEGG library. However, the appearance of anionic sugar lactones in the docking results strengthened this conviction.

Initially, eight lactones were synthesized from the monoacid derivatives of the D- and L-pentoses and then phosphorylated (Scheme 1). Of these eight compounds only D-lyxonolactone-5-phosphate (**1**) and L-ribonolactone-5-phosphate (**8**) proved to be substrates. The stereochemistry of the hydroxyl substituents at C2 and C3 of the two substrates are identical. However, the stereochemistry of the phosphorylated substituent at C4 is in either the R- or S-configuration for catalytic activity. Lmo2620 was also assayed with 2-deoxy-D-ribonolactone-5-phosphate (**9**) and L-ascorbate-6-phosphate (**10**) but no catalytic activity was found for either of these two compounds. The unphosphorylated sugar lactones and the diacid sugar lactones (compounds **11–22** of Scheme 2) were not substrates.

Bh0225 was also screened with the same set of potential substrates. No catalytic activity was detected with any of the organophosphates, unphosphorylated lactones or the diacid sugar lactones. Bh0225 preferentially hydrolyzes D-lyxonolactone-5-phosphate (**1**) and L-ribonolactone-5-phosphate (**8**) but none of the other phosphorylated lactones (compounds **2–7, 9**, and **10**). In addition to these compounds, we also synthesized D-mannono-1,5-lactone-6-phosphate but no hydrolytic activity could be detected by ^{31}P -NMR upon the addition of Bh0225. This compound was synthesized enzymatically using yeast hexokinase and MgATP to phosphorylate D-mannono-1,5-lactone. The kinetic parameters for the metal substituted forms of Lmo2620 and Bh0225 were determined and the results are presented in Table 4. [Co/Co]-Lmo2620 is the most active enzyme with values of k_{cat}/K_m of $\sim 10^5 \text{ M}^{-1} \text{ s}^{-1}$ for the hydrolysis of compounds **1** and **8**. [Co/Co]-Bh0225 is the most active form of this enzyme and the values of k_{cat}/K_m for the hydrolysis of compounds **1** and **8** exceed $10^5 \text{ M}^{-1} \text{ s}^{-1}$.

^{31}P NMR Analysis

NMR spectroscopy was used to examine the enzyme-catalyzed hydrolysis of compound **1** to confirm that the lactone ring was hydrolyzed rather than the monophosphate ester. The hydrolysis of compound **1** using [Co/Co]-Bh0225 was monitored as a function of time at pH 7.1 by ^{31}P NMR spectroscopy. The substrate has a single ^{31}P resonance at 4.03 ppm and the D-lyxonate-5-phosphate product has a single resonance at 3.72 ppm. The changes in the ^{31}P NMR spectra are shown in Figure 4. The resonance for the substrate at 4.03 ppm decreased and the resonance for the product at 3.72 ppm increased until all of the substrate was consumed. No inorganic phosphate or any other species was observed. These results demonstrated that Bh0225 is a lactonase and hydrolyzed compound **1** to D-lyxonate-5-phosphate.

Computational Docking to Active Site of Lmo2620

We explored the detailed binding determinants of the two substrates, compounds **1** and **8**, by docking their HEI structures. In the docked pose of compound **1** the phosphate group interacts with Arg-275, while both hydroxyl groups hydrogen bond with Lys-97. For compound **8**, on the other hand, the C2 hydroxyl interacts with Lys-97 while the C3 hydroxyl interacts with Tyr-71. As with compound **1**, the phosphate of **8** interacts with Arg-275 (Figures 6A and 6B). The poses of both compounds are positioned such that the reactive center of the HEI corresponds to an attack of the catalytic hydroxide from the *si*-face of the lactone ring.

DISCUSSION

Catalytic Functions in Cog1735

Enzymes within cog1735 previously have been shown to catalyze the hydrolysis of lactones and organophosphate triesters (17, 28–33, 35). Here, the substrate profiles for two previously uncharacterized proteins from group 5 of cog1735 (Lmo2620 and Bh0225) were determined through a collective strategy of bioinformatics, docking, three-dimensional structure determination, and library screening. At the initiation of this investigation, Lmo2620 was annotated as a phosphotriesterase family protein and Bh0225 was classified as a hypothetical protein by NCBI. No further functional information was available for any of the remaining proteins from group 5 of cog1735 (Figure 1).

Interrogation of Lmo2620 and Bh0225

The three-dimensional structure of Lmo2620 was determined to a resolution of 1.6 Å as a necessary first step for our structure-based approach to deciphering enzyme specificity. The overall fold of this protein is that of a distorted TIM-barrel and the active site is articulated by a binuclear metal center with a bridging carboxylated lysine that is a hallmark of the amidohydrolase superfamily (22). In Lmo2620 there are two additional phosphates bound at or near the active site. The proximal phosphate bridges the two divalent cations and displaces the hydroxide that usually coordinates the two metal ions (22). The distal phosphate is ~7.1 Å away from the proximal phosphate and forms polar interactions with Lys-242, Lys-244, Arg-275 and Tyr-278. These four residues are invariant within group 5 of cog1735. This observation, and the preponderance of lactone substrates for other proteins in cog1735, formed the primary structural basis for our initial prediction that Lmo2620 and Bh0225 would catalyze the hydrolysis of phosphorylated sugar lactones. Concurrently, the high-energy intermediate (HEI) form of KEGG had been docked against the X-ray structure of Lmo2620. The high enrichment of anionic lactones when the high-energy intermediate (HEI) form of KEGG was docked against the x-ray structure of Lmo2620 further strengthened our conviction. However, phosphorylated lactones were not observed in the initial docking calculations, owing to their absence from the library.

Substrate Identification

Two excellent substrates were discovered for Lmo2620 and Bh0225. These compounds, α -lyxonolactone-5-phosphate (**1**) and λ -ribonolactone-5-phosphate (**8**) have k_{cat}/K_m values of approximately $10^5 \text{ M}^{-1} \text{ s}^{-1}$. There is a requirement for a hydroxyl group at C2 with an S-configuration and a hydroxyl group at C3 with an R-configuration. The substituent at C4 can either be in the R- or S-configuration. Unfortunately, we were unable to synthesize comparable phosphorylated γ - or δ -lactones made from hexonic acids and thus it cannot be determined at this time whether these compounds would also function as substrates for these enzymes. However, we determined that the carboxylate mimics shown in Scheme 2 are not substrates for these two enzymes.

Model for Substrate Binding in Active Site

All of the structurally characterized members of the amidohydrolase superfamily catalyze the cleavage of esters and amides from the same relative face of the functional group that is hydrolyzed (22, 42). For esters of the type hydrolyzed by Lmo2620 it is the *si*-face of the lactone that is attacked by the bridging hydroxide (17). This stereochemical preference is recapitulated by the docking: the computed poses of compounds **1** and **8** feature an intermediate that corresponds to an attack to the *si*-face of the substrate. In this pose, the phosphate moiety at C5 is positioned such that it interacts with the same residues as the distal phosphate visible in the x-ray structure of Lmo2620. The hydroxyl groups interact

with Lys-97 and Tyr-71 for both substrates (Figure 6). At the same time, there is sufficient flexibility in the active site to enable the phosphate moiety of substrates with either stereochemistry at C4 to interact with the phosphate binding residues.

From a methodological standpoint, certain caveats bear mentioning. Perhaps most importantly, whereas the strategy of docking HEI forms of the KEGG metabolite library did identify what ultimately turned out to be the correct reactive center, the hydrolysis of sugar lactones, and even identified anionic forms of the lactones, it failed to identify what were ultimately the correct substrates. Whereas, in retrospect, sugar lactone phosphates do rank highly and would have been highly enriched in the docking screen, they were not in the original library. This highlights both the relative novelty of the activity discovered, and an ongoing challenge with docking a pre-existing metabolite library such as KEGG. As long as we depend upon substrates of enzymes of unknown activity being known, at full atomic detail, in KEGG, we will miss specific chemotypes. We cannot currently find substrates that are not already within KEGG by this approach. For many enzymes of unknown function this will likely be a serious liability, meriting further effort to overcome it.

Structure Comparison of Lmo2620 and Dr0930

Dr0930 (PDB: 3FDK and 3HTW) was previously identified from cog1735 as an enzyme that efficiently hydrolyzes simple δ - and γ -lactones with an alkyl substitution on the carbon adjacent to the ring oxygen. Lmo2620 and Dr0930 share ~30% sequence identity to each other and their structures are highly similar. The major difference between these two proteins is the sequence (Figure 7) and conformation of the loops that follow the eight β -strands. The four residues that are important for the recognition of the phosphate moiety in Lmo2620 (Lys-242, Lys-244, Arg-275 and Tyr-278) are replaced with four hydrophobic residues Leu-231, Met-234, Trp-267 and Leu-270 in Dr0930. These features are consistent with the observations that Lmo2620 hydrolyzes lactones with a phosphate moiety, while Dr0930 hydrolyzes lactones with a long hydrophobic tail. Conservation of these residues will be a significant clue in the ongoing assault on the substrate specificity for the remaining enzymes of unknown function in cog1735.

Possible Metabolic Role of Lmo2620 and Bh0225

The specific metabolic role of Lmo2620 and Bh0225 in their respective organisms has not been established. The two substrates for these enzymes are D -lyxonolactone-5-phosphate and L -ribonolactone-5-phosphate. The sugar precursors to these compounds, D -lyxose and L -ribose, are relatively rare sugars and not abundant in nature (50, 51). However, it has been reported that a mutant strain of *Escherichia coli* K12 can grow on D -lyxose as the sole carbon and energy source (52). It is also known that *Cohnella laevoribosii* RI-39 can grow in a defined medium with L -ribose as the sole carbon source (53). We speculate that Lmo2620 and Bh0225 may be involved in a pentose (or hexose) phosphate pathway that is initiated by the phosphorylation of the lactone followed by the hydrolysis of the lactone to the acid sugar phosphate. Acid sugars of this type can be dehydrated to a keto-sugar and subsequently cleaved via an aldolase or ketolase.

Strategy for Functional Annotation

Bioinformatics, three-dimensional structure determination, computational docking, and library screening were used synergistically to successfully identify two novel substrates for enzymes of unknown function in cog1735. The identification of two phosphates in the active site of Lmo2620 was leveraged with the knowledge that most of the functionally annotated enzymes within cog1735 hydrolyze lactone substrates. The preponderance of lactones in the docking calculations led to the synthesis of phosphorylated lactones and the identification of two excellent substrates for Lmo2620 and Bh0225. Based upon the conserved residues in

the active site of Lmo2620 that appear critical for substrate recognition (Tyr-71, Lys-97, Lys-242, Lys-244, Arg-275, Tyr-278) we now predict that additional members of group **5** of cog1735 will have the same substrate profile. These proteins are listed in Table S2 of the Supporting Information.

It is also possible to predict some of the substrate attributes for enzymes of unknown function in groups **1** and **6** of cog1735 (Figure 1). MS53_0025 from *Mycoplasma synoviae* 53 (gi|71894050), a representative from group **6**, has the five conserved residues (except for two lysine/arginine substitutions) from group **5** that appear to form the phosphate binding site. This strongly suggests that this enzyme will also utilize a phosphorylated substrate. For the phosphotriesterase homology protein (PHP) from group **1**, the situation is less clear since there are multiple substitutions in the phosphate binding site (Figure 7). Nevertheless, this information provides significant leverage for the successful annotation of the remaining proteins in cog1735.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Siddhesh Kamat for the enzymatic synthesis of D-mannono-1,5-lactone 6-phosphate.

Funding

This work was supported in part by the NIH (GM 071790) and the Robert A. Welch Foundation (A-840).

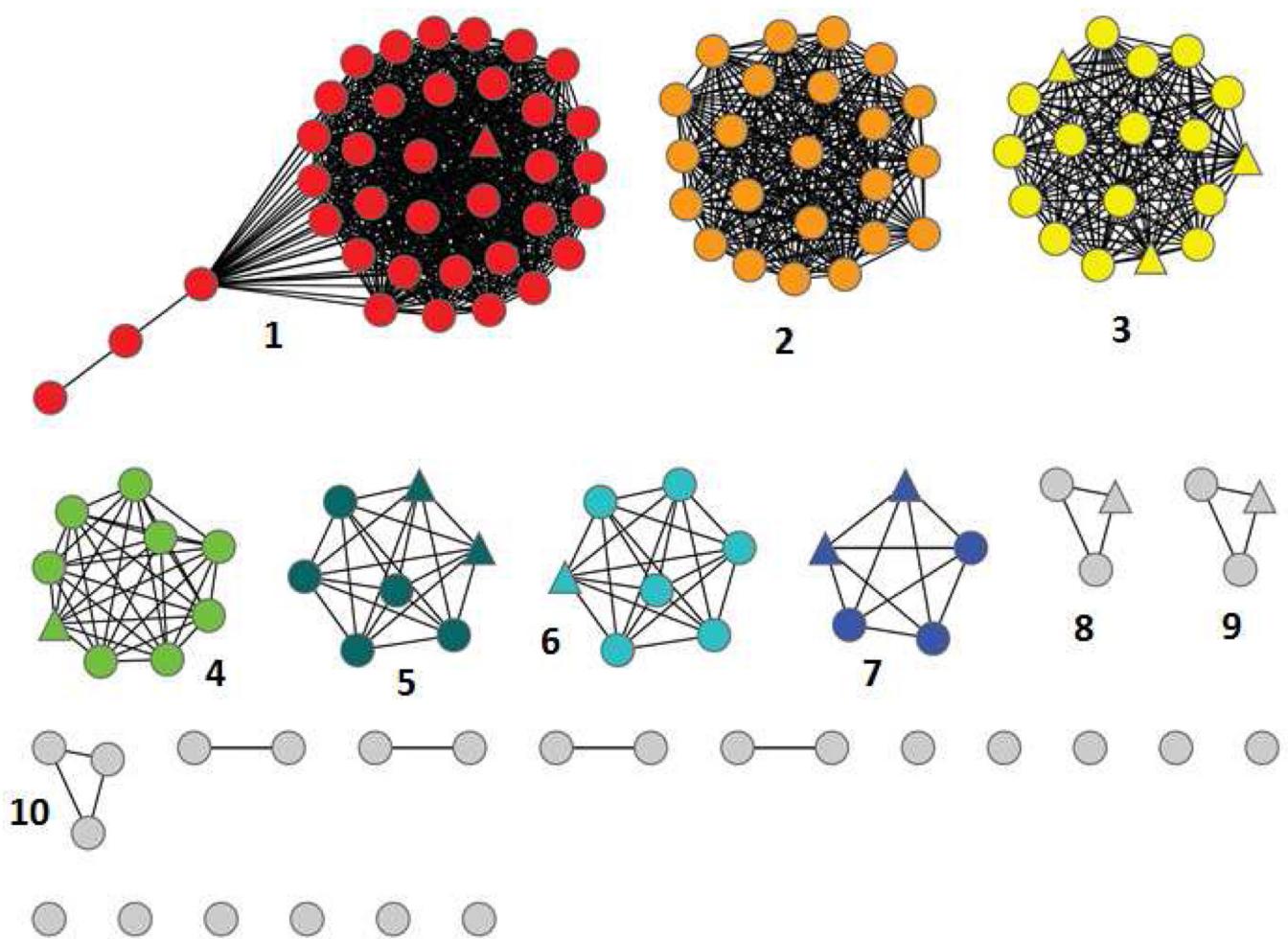
References

1. Juncker AS, Jensen LJ, Pierleoni A, Bernsel A, Tress ML, Bork P, von Heijne G, Valencia A, Ouzounis CA, Casadio R, Brunak S. Sequence-based feature prediction and annotation of proteins. *Genome Biol.* 2009; 10:206.1–206.6. [PubMed: 19226438]
2. Pieper U, Chiang R, Seffernick JJ, Brown SD, Glasner ME, Kelly L, Eswar N, Sauder JM, Bonanno JB, Swaminathan S, Burley SK, Zheng X, Chance MR, Almo SC, Gerlt JA, Raushel FM, Jacobson MP, Babbitt PC, Sali A. Target selection and annotation for the structural genomics of the amidohydrolase and enolase superfamily. *J. Struct. Funct. Genomics.* 2009; 10:107–125. [PubMed: 19219566]
3. Wu CH, Aweiler R, Bairoch A, Natale DA, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Mazumander R, O'Donova C, Redaschi N, Suzek B. The universal protein resource (Uniprot): an expanding universe of protein information. *Nucleic Acids Res.* 2006; 34:D187–D191. [PubMed: 16381842]
4. Graillie M, Baltaze J-P, Leulliot N, Liger D, Quevillo-Cheruel S, von Tibeburgh H. Structure-based functional annotation: yeast ymr099c codes for a D-hexose-6-phosphate mutarotase. *J. Biol. Chem.* 2006; 281:30175–30185. [PubMed: 16857670]
5. Friedberg I, Jambon M, Godzik A. New avenues in protein function prediction. *Protein Sci.* 2006; 15:1527–1529. [PubMed: 16731984]
6. Hall RS, Agarwal R, Hitchcock D, Sauder JM, Burley SK, Swaminathan S, Raushel FM. Discovery and structure determination of the orphan enzyme isoxanthopterin deaminase. *Biochemistry.* 2010; 49:4374–4382. [PubMed: 20415463]
7. Xiang DF, Patskovsky Y, Xu C, Fedorov AA, Fedorov EV, Sicco AA, Sauder JM, Burley SK, Almo SC, Raushel FM. Functional identification and structure determination of two novel prolidases from cog 1228 in the amidohydrolase superfamily. *Biochemistry.* 2010; 49:6791–6803. [PubMed: 20604542]
8. Cummings JA, Nguyen TT, Fedorov AA, Kolb P, Xu C, Fedorov EV, Shoichet BK, Barondeau DP, Almo SC, Raushel FM. Structure, mechanism, and substrate profile for Sco3058: the closest

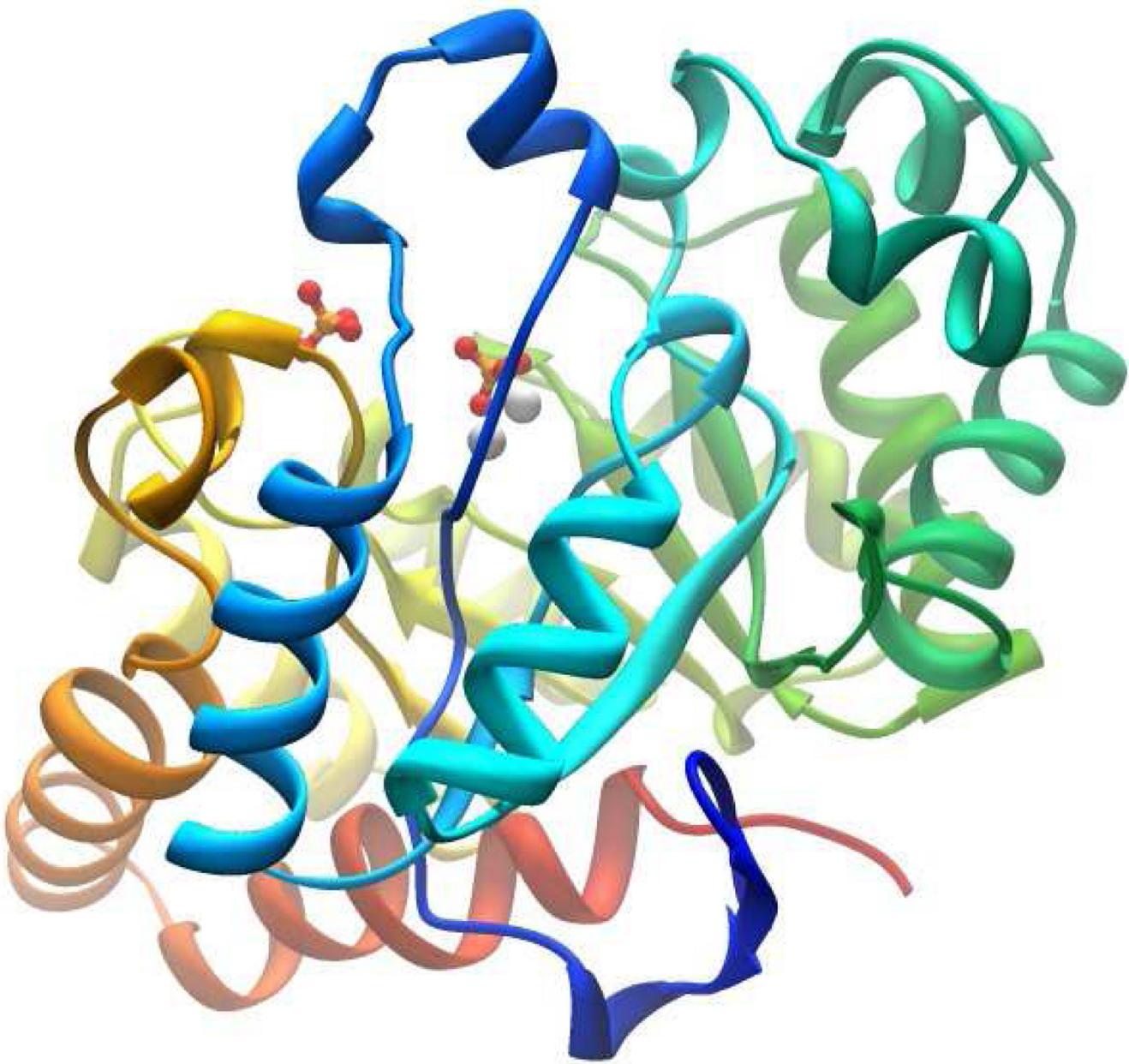
- bacterial homologue to human renal dipeptidase. *Biochemistry*. 2010; 49:611–622. [PubMed: 20000809]
9. Cummings JA, Fedorov AA, Xu C, Brown S, Fedorov E, Babbitt PC, Almo SC, Raushel FM. Annotating enzymes of uncertain function: The deacylation of D-amino acids by members of the amidohydrolase superfamily. *Biochemistry*. 2009; 48:6469–6481. [PubMed: 19518059]
10. Xiang DF, Patskovsky Y, Xu C, Meyer AJ, Sauder JM, Burley SK, Almo SC, Raushel FM. Functional identification of incorrectly annotated prolidases from the amidohydrolase superfamily of enzymes. *Biochemistry*. 2009; 48:3730–3742. [PubMed: 19281183]
11. Nguyen TT, Brown S, Fedorov AA, Fedorov EV, Babbitt PC, Almo SC, Raushel FM. At the periphery of the amidohydrolase superfamily: Bh0493 from *Bacillus halodurans* catalyzes the isomerization of d-galacturonate to d-tagaturonate. *Biochemistry*. 2008; 47:1194–1206. [PubMed: 18171028]
12. Marti-Arbona R, Xu C, Steele S, Weeks A, Kuty GF, Seibert CM, Raushel FM. Annotating enzymes of unknown function: N-Formimino-l-glutamate deiminase is a member of the amidohydrolase superfamily. *Biochemistry*. 2006; 45:1997–2005. [PubMed: 16475788]
13. Hermann JC, Marti-Arbona R, Fedorov AA, Fedorov E, Almo SC, Shoichet BK, Raushel FM. Structure based activity prediction for an enzyme of unknown function. *Nature*. 2007; 448:775–779. [PubMed: 17603473]
14. Hermann JC, Ghanem E, Li Y, Raushel FM, Irwin JJ, Shoichet BK. Predicting substrates by docking high-energy intermediates to enzyme structures. *J. Am. Chem. Soc.* 2006; 128:15882–15891. [PubMed: 17147401]
15. Irwin JJ, Raushel FM, Shoichet BK. Virtual screening against metalloenzymes for inhibitors and substrates. *Biochemistry*. 2005; 44:12316–12328. [PubMed: 16156645]
16. Xiang DF, Xu C, Kumaran D, Brown AC, Sauder JM, Burley SK, Swaminathan S, Raushel FM. Functional annotation of two new carboxypeptidases from the amidohydrolase superfamily of enzymes. *Biochemistry*. 2009; 48:4567–4576. [PubMed: 19358546]
17. Xiang DF, Kolb P, Fedorov AA, Meier MM, Fedorov LV, Nguyen TT, Sterner R, Almo SC, Shoichet BK, Raushel FM. Functional annotation and three-dimensional structure of Dr0930 from *Deinococcus radiodurans*, a close relative of phosphotriesterase in the amidohydrolase superfamily. *Biochemistry*. 2009; 48:2237–2247. [PubMed: 19159332]
18. Kamat SS, Fan H, Sauder JM, Burley SK, Shoichet BK, Sali A, Raushel FM. Enzymatic deamination of the epigenetic base N-6-methyladenine. *J. Am. Chem. Soc.* 2011; 133:2080–2083. [PubMed: 21275375]
19. Hitchcock DS, Fedorov AA, Fedorov EV, Dangott LJ, Almo SC, Raushel FM. Rescue of the orphan enzyme isoguanine deaminase. *Biochemistry*. 2011; 50:5555–5557. [PubMed: 21604715]
20. Goble AM, Zhang Z, Sauder JM, Burley SK, Swaminathan S, Raushel FM. Pa0148 from *Pseudomonas aeruginosa* catalyzes the deamination of adenine. *Biochemistry*. 2011; 50:6589–6597. [PubMed: 21710971]
21. Holm L, Sander C. An evolutionary treasure: unification of a broad set of amidohydrolase related to urease. *Proteins*. 1997; 28:72–82. [PubMed: 9144792]
22. Seibert CM, Raushel FM. Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry*. 2005; 44:6383–6391. [PubMed: 15850372]
23. Martynowski D, Eyobo Y, Li T, Yang K, Liu A, Zhang H. Crystal structure of alpha-amino-beta-carboxymuconate-epsilon-semialdehyde decarboxylase: Insight into the active site and catalytic mechanism of a novel decarboxylation reaction. *Biochemistry*. 2006; 45:10412–10421. [PubMed: 16939194]
24. Liu A, Zhang H. Transition metal-catalyzed nonoxidative decarboxylation reactions. *Biochemistry*. 2006; 45:10407–10411. [PubMed: 16939193]
25. Hara H, Masai E, Katayama Y, Fukuda M. The 4-oxalomesaconate hydratase gene, involved in the protocatechuate 4,5-cleavage pathway, is essential to vanillate and syringate degradation in *Sphingomonas paucimobilis* SYK-6. *J. Bacteriol.* 2000; 182:6950–6957. [PubMed: 11092855]
26. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein function and evolution. *Nucleic Acids*. 2000; 28:33–36.

27. Atkinson HJ, Morris JH, Ferrin TE, Babbitt PC. Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS One*. 2009; 4:e4345. [PubMed: 19190775]
28. Ghanem E, Raushel FM. Detoxification of organophosphate nerve agents by bacterial phosphotriesterase. *Toxicol. Appl. Pharmacol.* 2005; 207:459–470. [PubMed: 15982683]
29. Merone L, Mandrich L, Rossi M, Manco G. A thermostable phosphotriesterase from the archeon *Sulfolobus solfataricus*: cloning, overexpression and properties. *Extremophiles*. 2005; 9:297–305. [PubMed: 15909078]
30. Afriat L, Roodveldt C, Manco G, Tawfik DS. The latent promiscuity of newly identified microbial lactonases is linked to a recently diverged phosphotriesterase. *Biochemistry*. 2005; 45:13677–13686. [PubMed: 17105187]
31. Uroz S, Oger PM, Chapelle E, Adeline M-T, Faure D, Dessaux Y. A *Rhodococcus qsdA*-encoded enzyme defines a novel class of large-spectrum quorum-quenching lactonases. *Appl. Environ. Microbiol.* 2008; 74:1357–1366. [PubMed: 18192419]
32. Chow JY, Wu L, Yew WS. Direct evolution of a quorum-quenching lactonase from *Mycobacterium avium* subsp. *Paratuberculosis* K-10 in the amidohydrolase superfamily. *Biochemistry*. 2009; 48:4344–4553. [PubMed: 19374350]
33. Hawwa R, Aikens J, Turner RJ, Santarsiero BD, Mesecar AD. Structural basis for thermostability revealed through the identification and characterization of a highly thermostable phosphotriesterase-like lactonase from *Geobacillus stearothermophilus*. *Arch. Biochem. Biophys.* 2009; 488:109–120.
34. Buchbinder JL, Stephenson RC, Dresser MJ, Pitera JW, Scanlan TS, Fletterick RJ. Biochemical characterization and crystallographic structure of an *Escherichia coli* protein from the phosphotriesterase gene family. *Biochemistry*. 1998; 37:5096–5106. [PubMed: 9548740]
35. Hawwa R, Larsen SD, Ratia K, Mesecar AD. Structure-based and random mutagenesis approaches increase the organophosphate-degrading activity of a phosphotriesterase homologue from *Deinococcus radiodurans*. *J. Mol. Biol.* 2009; 393:36–57. [PubMed: 19631223]
36. Bennis K, Calinaud P, Gelas J, Ghobi M. A new route to some enantiomerically pure substituted morpholines from d-ribono- and d-gulono-1,4-lactones. *Carbohydrate Research*. 1994; 264:33–44.
37. Humphlett WJ. Synthesis of some esters and lactone of aldonic acids. *Carbohydrate Research*. 1967; 4:157–164.
38. Liao M-L, Wang X-Y, Chung C, Liang Y-T, Seib PA. Synthesis of l-ascorbate 6-phosphate. *Carbohydrate Research*. 1998; 176:73–77.
39. Bock K, Lundt I, Pedersen C. Preparation of some bromodeoxyaldonic acids. *Carbohydrate Research*. 1979; 68:313–319.
40. Niemann C, Link KP. Synthesis of the hexuronic acids I. The synthesis of *dl*- galacturonic acid from mucic acid. *J. Biol. Chem.* 1932; 95:203–211.
41. Bose RJ, Hullar TL, Lewis BA, Smith F. Isolation of the 1,4-and 6,3-lactones of d-glucaric acid. *J. Org. Chem.* 1961; 26:1300–1301.
42. Kamat SS, Bagaria A, Kumaran D, Holmes-Hampton GP, Fan H, Sali A, Sauder JM, Burley SK, Lindahl PA, Swaminathan S, Raushel FM. Catalytic mechanism and three-dimensional structure of adenine deaminase. *Biochemistry*. 2011; 50:1917–1927. [PubMed: 21247091]
43. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. In: Carter, CWJ.; Sweet, RM.; Abelson, JN.; Simon, MI., editors. *Method. Enzymol.* Vol. 276. New York: Academic Press; 1997. p. 307-326.
44. Long F, Vagin A, Young P, Murshudov GN. BALBES: a Molecular Replacement Pipeline. *Acta Crystallogr.* 2008; D64:125–132.
45. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr.* 2004; D60:2126–2132.
46. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr.* 2010; D66:213–221.

- NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript
47. Lamzin VS, Wilson KS. Automated refinement of protein models. *Acta Crystallogr.* 1993; D49:129–147.
 48. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 2000; 28:27–30. [PubMed: 10592173]
 49. Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res.* 2006; 34:D354–D357. [PubMed: 16381885]
 50. Yeom S-J, Ji J-H, Kim N-H, Park C-S, Oh D-K. Substrate specificity of a mannose-6-phosphate isomerase from *Bacillus subtilis* and its application in the production of L-ribose. *App. Environ. Microbiol.* 2009; 75:4705–4710.
 51. Kwon H-J, Yeon S-J, Park C-S, Oh D-K. Substrate specificity of a recombinant d-lyxose isomerase from *Providencia stuartii* for monosaccharides. *J. Biosci. Bioeng.* 2010; 110:26–31. [PubMed: 20541111]
 52. Stevens FJ, Wu TT. Growth on d-lyxose of a mutant strain of *Escherichia coli* K-12 using a novel isomerase and enzymes related to d-xylose metabolism. *J. Gen. Microbiol.* 1976; 97:257–265. [PubMed: 796410]
 53. Cho E-A, Lee D-W, Cha Y-H, Lee S-J, Jung H-C, Pan J-G, Pyun Y-R. Characterization of a novel d-lyxose isomerase from *Cohnella laevoribosii* RI-39 sp. nov. *J. Bacteriol.* 2007; 189:1655–1663. [PubMed: 17189362]

**Figure 1.**

Cytoscape network representation (www.cytoscape.org) of the sequence relationships in cog1735 from the amidohydrolase superfamily. Each node in the network represents a single sequence and each edge (depicted as lines) represents the pairwise connection between two sequences with the most significant BLAST E-value (better than 1×10^{-80}). Lengths of edges are not meaningful except that sequences in tightly clustered groups are relatively more similar to each other than sequences with few connections. The triangular nodes represent proteins that have been functionally and/or structurally characterized. Additional information is provided in the text.

**Figure 2.**

Ribbon representation for the three-dimensional structure of [Zn/Zn]-Lmo2620 with two phosphates in the active site.

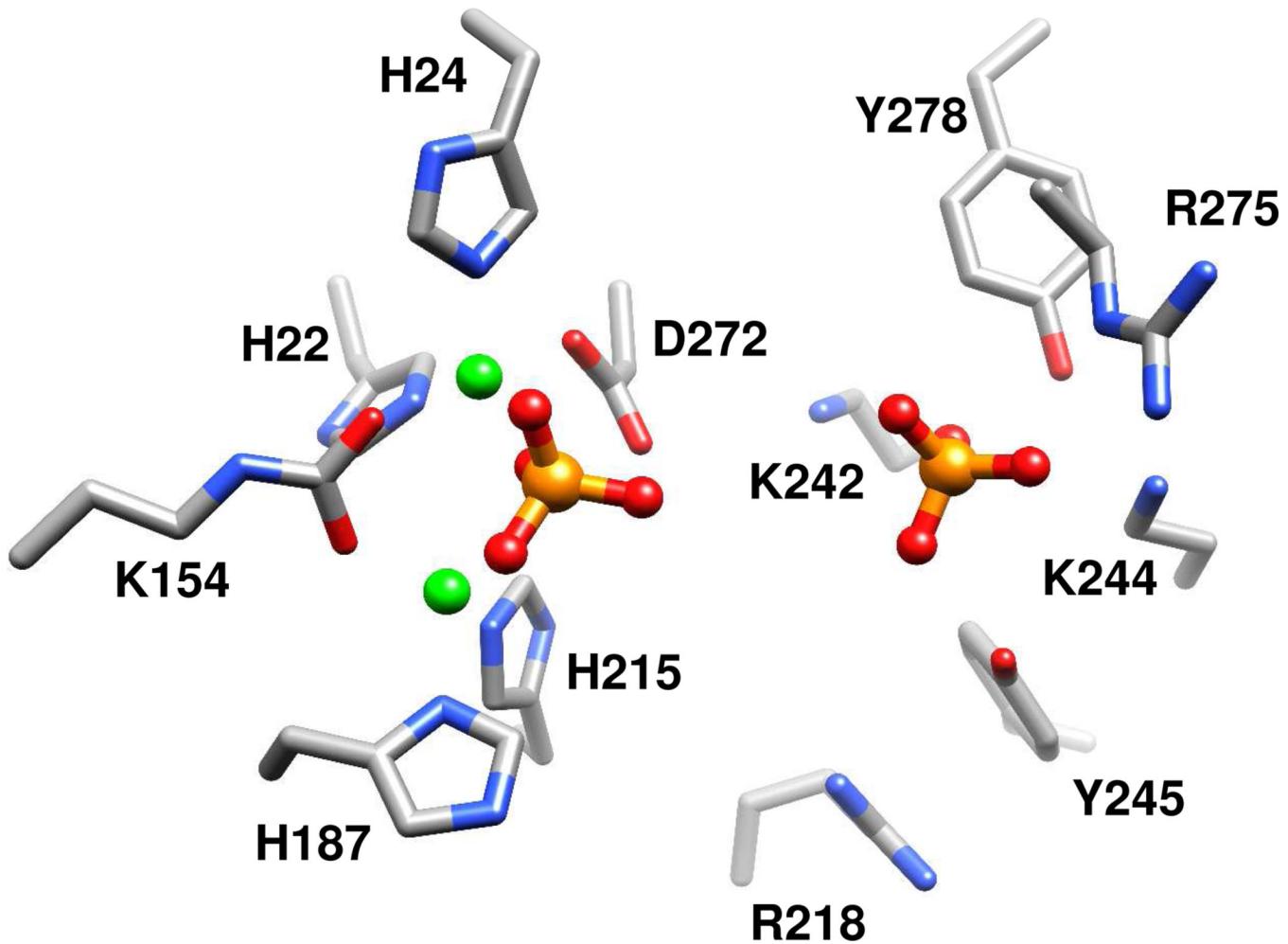


Figure 3.

Active site of [Zn/Zn]-Lmo2620 with the two bound phosphates. The two divalent zinc ions are presented as green spheres.

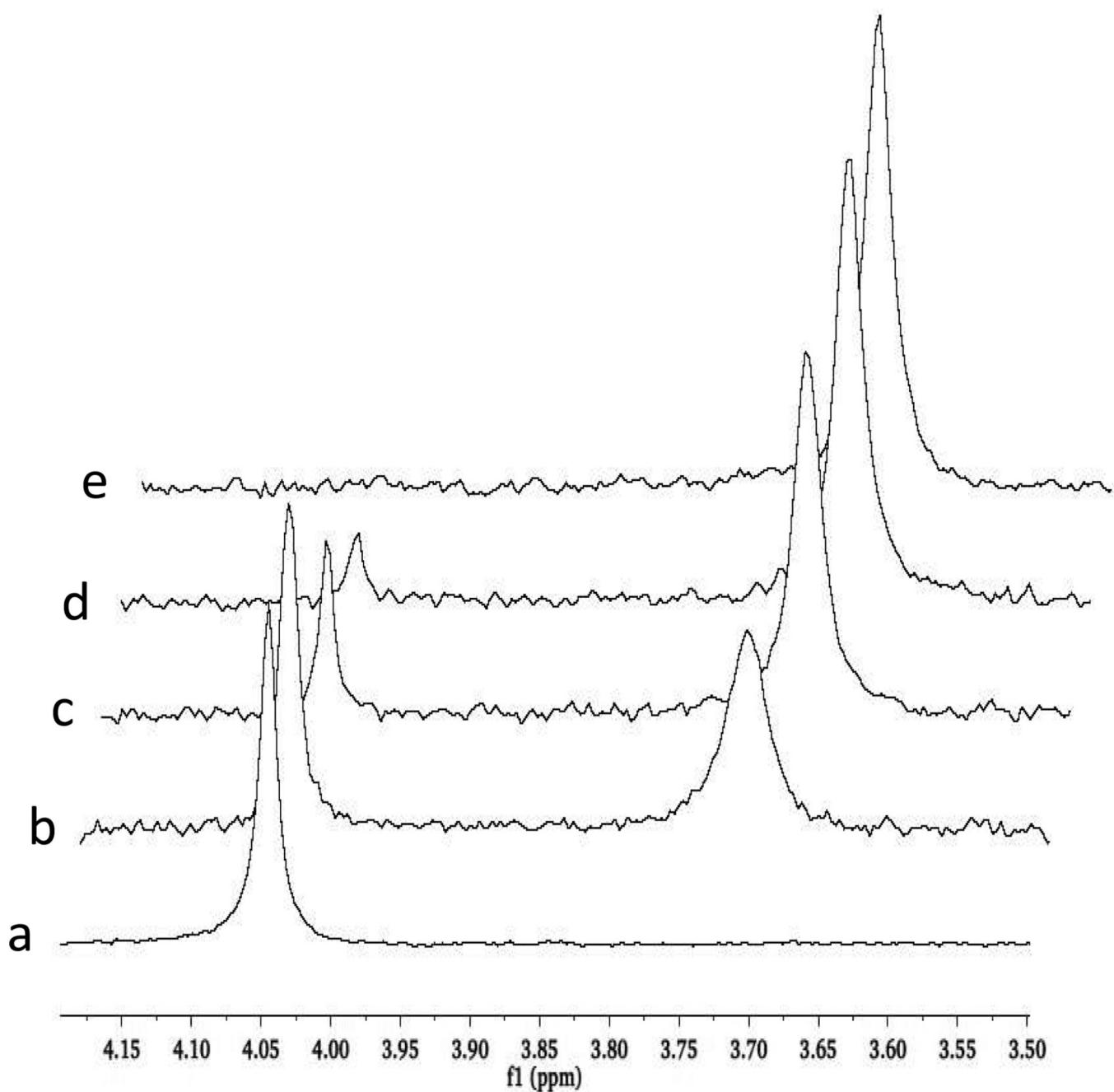
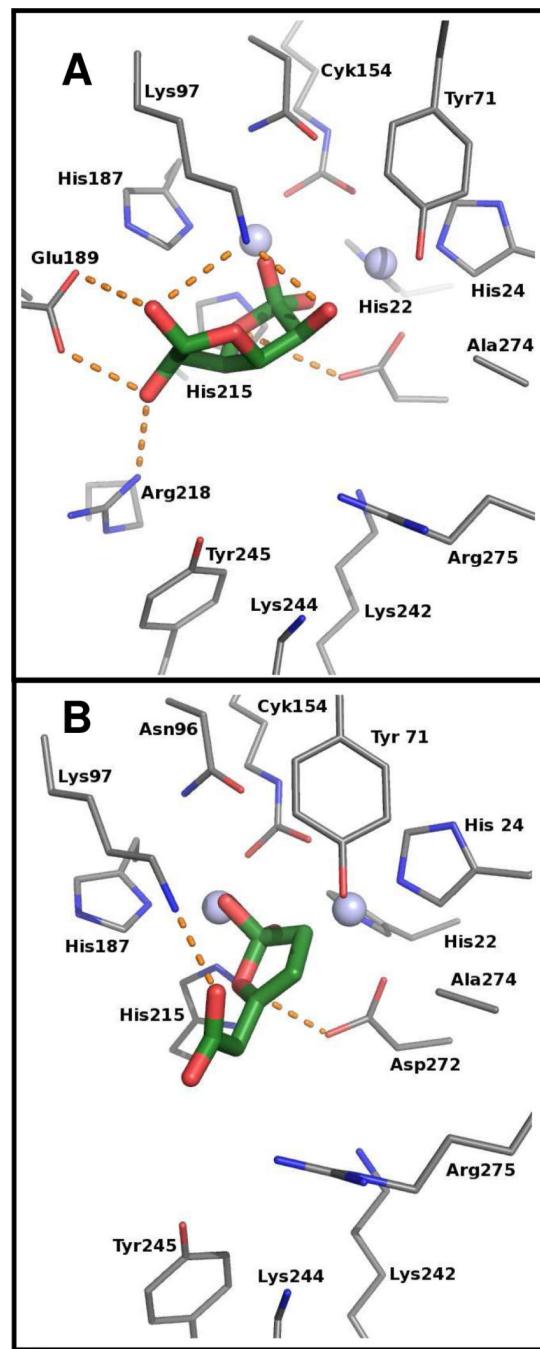
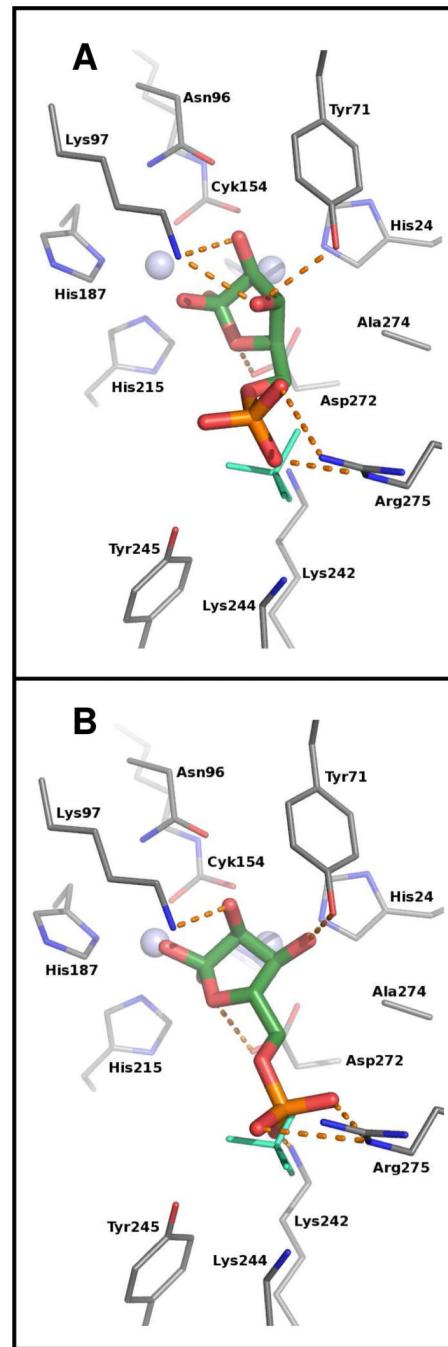


Figure 4.

Time dependent hydrolysis of D-lyxonolactone-5-phosphate catalyzed by Bh0225 followed by ^{31}P -NMR spectroscopy at pH 7.1 and 5 °C. (A), 0 minutes; (B), 5 minutes; (C), 20 minutes; (D), 35 minutes; (E), 60 minutes.

**Figure 5.**

Docked poses of characteristic high-ranking candidate substrates from the docking screen of the high-energy intermediate form of the KEGG library. (A) A sugar lactone, compound K5 in Table 3, ranked 217 in the docking screen. (B) A lactone carboxylate, ranked 91 in the docking screen. The high energy intermediate form of both compounds was docked and is represented; this group interacts with the catalytic metals. Key recognition residues are labeled. Putative hydrogen bonds are represented as amber dashed lines. Protein carbons are gray, ligand carbons green, oxygens are red nitrogens are blue.

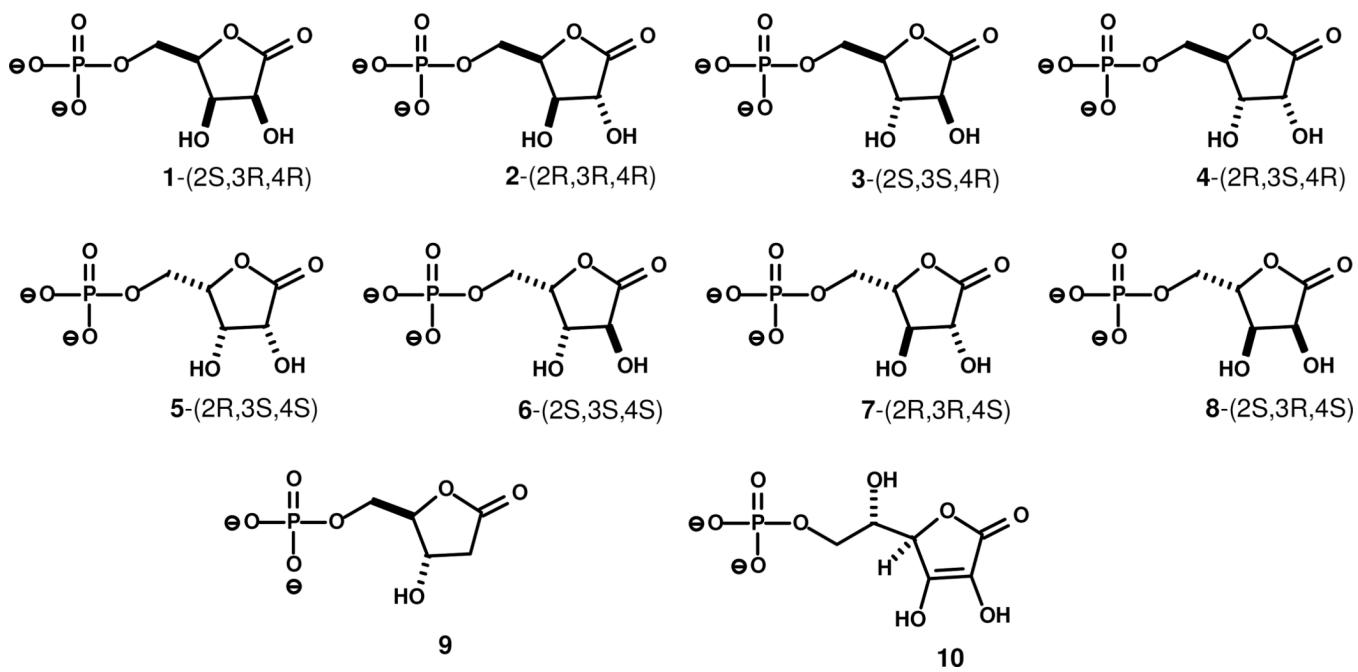
**Figure 6.**

Docked poses of the substrates with Lmo2620. The position of the phosphate ion from the original structure is shown in cyan; the protein and ligands are otherwise colored as in Figure 5. (A) Compound 1. (B) Compound 8.

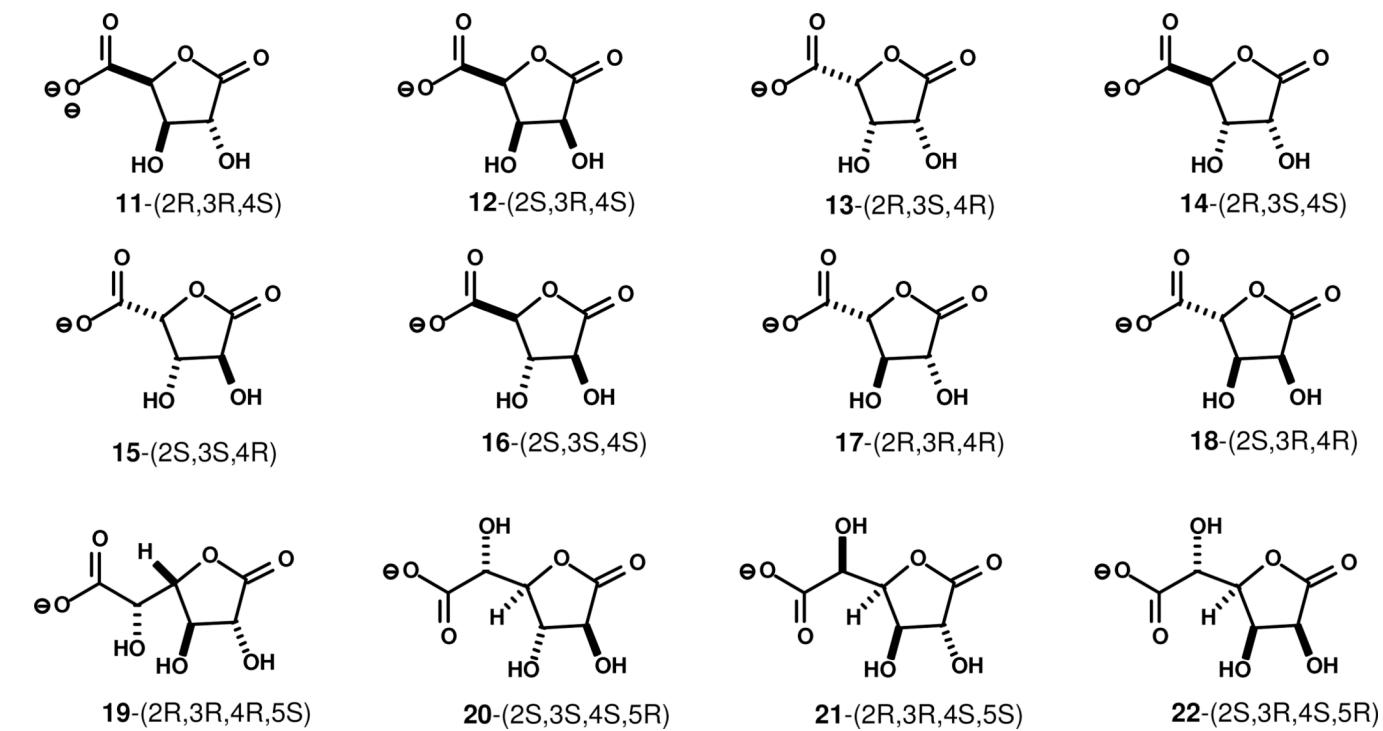
Lmo2620	--MSFIRTFYGDIAPEQLGFTYS H E H IVCVPAYWQERDAD--DLLLDDKEKSQLDVQDFA
Bh0225	--MSFIRTLRGDINP E ELGFTYS H E H IVCRPDYWVQRQEG--DLLLDDKEKS K RDV E DFK
MS53_0025	MENKFARTVLGDI P VEKL G ITDC H D H FIKNGGPEVEEHID--FLMLN-VDASIKEFKEFI
PHP	-----MSFDPTGYTLA H E H LHIDL S GFKN--V--DCRLDQYAFICQEMNDLM
Dr0930	---MTAQTVTGAVAAAQLGATLP H E H IVFGYPGYAGDVTLGPFDHAAALASCTETARALL
 Lmo2620	 DLGGKTIVDATAVD Y G RRVLDVAQIS--KETGIQIVGTAGFN K S FLWDGKIKPELKPII
Bh0225	RHGGNSIVDATAVD Y G RDVGAVKAIS--DELDVHIIGTAGFN K S FLWNAKIKEELKPII
MS53_0025	DRGGSTIVTMDPPN V G RDVLKLTLEIANAVKNLGGNVIMSTGFH K A KFY-----
PHP	TRGVRNVIEMTNRY M G RNAQFMLDVMR--ETGINVVACTGYY Q D AFFP-----
Dr0930	ARGIQTVVDA T P NDC G RNP <small>A</small> F <small>L</small> REVS--EATGLQILCATGFY Y E GEGAT-----
 Lmo2620	 G-DFETYYEWIENT--TTDKLTEFVVNEVNGLEGTPY-----KAGQV K F GTGYN
Bh0225	GGQYDTFAEWIDGS--TIHDLTEFVVREVEEGLEGTPW-----KAGQV K F GTGYN
MS53_0025	----DKYSSWLAVV--PTEEIVKMCVAEIEEGMDEYNYNGPVVKRSKAKAGII I K AGTGYG
PHP	-----EHVATR--SVQELAQEMVDEIEQGIDGTELK-----AGII E IGTSEG
Dr0930	-----TYFKFRASLGDAESEIYEMMRTEVTEGIAGTGIR-----AGVI K ASSRD
 Lmo2620	 MITPLEEK T IRAVARAH H ETKAPI H STEAGTM A LEQIEILK Q ENIP L YLS I G H M D R NL
Bh0225	RISPLEEK T IRAVARAH H ETKAPI H TEVG T MALEQ I ELL K SE G IN L EH V S I G H M D R NP
MS53_0025	AIDRLELKALEVAARTSILTGCPI L V H TQLGTMALEVAKH L IGFGANPD K IQ I SHLN K NP
PHP	KITPLEEKVFIAAALAHNQTGRPI S T H TSFSTM G LEQL A LLQAHGV D LSRV T V G H C D L KD
Dr0930	AITPYEQLF F RAA A RV Q RET G VPI I H TQEGQQGPQQ A LLTSLGADPARIM I G H M D G NT
 Lmo2620	 DPYYHKQVAKT-GAFMSFDGI---A K I K YAPESARIAA I LYLVSEG F E D Q I L VSG D T A R K
Bh0225	DPYYHEQIAKT-GAFLSF D G I---G K I K YAPESTRIHCILELVKKGYD Q Q I L IS G D T A R K
MS53_0025	DKYYYEKVIKETGV T LCFDGP---D R V K Y PD S LLA E NI K YLV D K Q K H IT L S D A R I
PHP	NLDNILKMIDL-GAYVQFD T I---G K N S Y PD E KRI A MLHAL R DRG L LN R VML S M D I T R R
Dr0930	DPAYHRETLRH-GVSIAFDRIGL Q G M V G T PTDAERLSV L T L L GE G YAD R LL L S H D S I W
 Lmo2620	 T Y Y K---HYGH P GP-LEYIAKKW V PRFIDEANE K G F D G E K L V K F V D N P A R C F T F K K
Bh0225	 T Y Y K---HYDYGLG-LEYIIAKW I P RF I SEANHRGF D G E A L I H T F F V E N P A R C F A F K L
MS53_0025	 L Y Q R---NYGLTK G K Q T F G L A Y L F D R F L P L K Q V G V S -KEA I F D I L V N N P K R V L A F E K
PHP	 S H L K---ANGG-----YGYDYL T T F I P Q L R Q S G F S -QADVD V M L R E N P S Q F F Q ---
Dr0930	 W L G R P P A I P A Al P A V K D W H P L H I S D D I L P D L R R G I T -EEQVG Q M T V G N P A R L F G ---
 Lmo2620	 -----
Bh0225	 -----
MS53_0025	RNF D P L K V S K E V L E K K E L N L N
PHP	 -----
Dr0930	 -----

Figure 7.

Amino acid sequence alignment of selected proteins from cog1735. Lmo2620 (group **5**), Bh0225 (group **5**), Dr0930 (group **7**), phosphotriesterase homology protein (group **1**) from *E. coli* (locus tag: b3379) and MS53_0025 from *Mycoplasma synoviae* 53 (group **6**). The metal binding ligands are shown in red. The residues that have been shown to facilitate the binding of substrates in the active site of Lmo2620 and Bh0225 are highlighted in green. The corresponding residues in the other three proteins are highlighted in gray.



Scheme 1.



Scheme 2.

Table 1

Data Collection and Refinement Statistics for Lmo2620

<i>Data collection</i>	
Beamline	NSLS X4A
wavelength (Å)	0.97915
space group	P2 ₁
# of molecules in a.u.	6
<i>Unit cell parameters</i>	
a (Å)	152.72
b (Å)	62.93
c (Å)	152.74
β (°)	90.01
resolution (Å) ^a	40–1.6 (1.66–1.60)
unique reflections ^a	360927 (24807)
completeness (%) ^a	94.1 (65.0)
R _{merge} ^a	0.087 (0.419)
Average I/σ ^a	18.7 (3.9)
<i>Refinement</i>	
resolution	40.0–1.6
R _{cryst}	0.178
R _{free}	0.197
rmsd for bonds (Å)	0.006
rmsd for angles (°)	1.06
# of protein atoms	15842
# of waters	1493
# of Zn ²⁺ ions	12
# of phosphate ions	12
PDB entry	3PNZ

^aNumbers in parentheses indicate values for the highest resolution shell.

Metal Content of Lmo2620 and Bh0225

Table 2

Enzyme	Zn	Mn	Co	Fe
[Zn/Zn]-Lmo2620	1.9	-	-	0.2
[Mn/Mn]-Lmo2620	0.2	1.8	-	0.1
[Co/Co]-Lmo2620	0.2	-	1.7	0.2
[Zn/Zn]-Bh0225	1.8	-	-	0.1
[Mn/Mn]-Bh0225	0.2	1.6	-	0.2
[Co/Co]-Bh0225	0.1	-	2.0	-

Table 3

Results of docking the HEI version of KEGG to Lmo2620.

Reaction type	Representative structures	N _{comp} ^a	ranks ^b	N _{int} ^c
lactam opening	 K1, rank 183	54	2–497	4
acyclic deamination	 K2, rank 3	26	3–487	1
ester cleavage	 K3, rank 9	37	9–488	8
amide cleavage	 K4, rank 16	103	16–500	13
lactone opening	 K5, rank 217	46	34–491	34
exocyclic deamination	 K6, rank 278	5	278–357	0

Table of enriched chemotypes and representative structures from among the top 500 molecules of the KEGG HEI database ranked according to DOCK score.

^aNumber of molecules of a particular chemotype with a catalytically competent docking pose among the top 500.

^bHighest and lowest rank among the top 500.

^cNumber of molecules in catalytically competent poses that have all of their polar interactions satisfied.

Table 4

Kinetic parameters for Lmo2620 and Bh0225.

Substrate	Enzyme	$k_{cat}(s^{-1})$	$k_m(mM)$	$k_{cat}/k_m(M^{-1}s^{-1})$
1	[Zn/Zn]-Lmo2620	1.1 ± 0.1	0.46 ± 0.05	(2.4 ± 0.3) × 10 ³
	[Mn/Mn]-Lmo2620	1.7 ± 0.1	0.23 ± 0.02	(7.4 ± 0.7) × 10 ³
	[Co/Co]-Lmo2620	15 ± 1	0.25 ± 0.03	(6.1 ± 0.7) × 10 ⁴
	[Zn/Zn]-Bh0225	1.8 ± 0.05	0.40 ± 0.04	(4.5 ± 0.5) × 10 ³
	[Mn/Mn]-Bh0225	1.6 ± 0.04	0.20 ± 0.02	(8.0 ± 0.8) × 10 ³
	[Co/Co]-Bh0225	24 ± 1	0.24 ± 0.02	(1.0 ± 0.1) × 10 ⁵
8	[Zn/Zn]-Lmo2620	1.5 ± 0.1	0.43 ± 0.03	(3.5 ± 0.3) × 10 ³
	[Mn/Mn]-Lmo2620	3.6 ± 0.1	0.45 ± 0.04	(8.0 ± 0.7) × 10 ³
	[Co/Co]-Lmo2620	18 ± 1	0.24 ± 0.03	(7.3 ± 0.6) × 10 ⁴
	[Zn/Zn]-Bh0225	1.6 ± 0.1	0.32 ± 0.02	(5.0 ± 0.5) × 10 ³
	[Mn/Mn]-Bh0225	1.8 ± 0.1	0.21 ± 0.02	(8.6 ± 0.9) × 10 ³
	[Co/Co]-Bh0225	30 ± 1	0.2 ± 0.02	(1.5 ± 0.15) × 10 ⁵