Published in final edited form as:
Biochemistry. 2007 November 27; 46(47): 13606–13615. doi:10.1021/bi701439j.

Gene identification and structural characterization of the PLP degradative protein 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase from *Mesorhizobium loti MAFF303099*

Tathagata Mukherjee[#], **Kathryn M. McCulloch**[#], **Steven E. Ealick**[*], and **Tadhg P. Begley**[*]
Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853

Abstract

The function of the mlr6791 gene from *Mesorhizobium loti* MAFF303099 has been identified. This gene encodes 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase (HMPDdc), an enzyme involved in the catabolism of pyridoxal 5-phosphate (Vitamin B_6). This enzyme was overexpressed in *Escherichia coli* and characterized. HMPDdc is a 26 kDa protein that catalyzes the decarboxylation of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate to 3-hydroxy-2-methylpyridine-5-carboxylate. The K_M and k_{cat} were found to be 366 μM and 0.6 s^{-1} respectively. The structure of this enzyme was determined at 1.9 Å resolution using SAD phasing and belongs to the class II aldolase/adducin superfamily. While the decarboxylation of hydroxy-substituted benzene rings is a common motif in biosynthesis, the mechanism of this reaction is still poorly characterized. The structural studies described here suggest that catalysis of such decarboxylations proceeds by an aldolase-like mechanism.

In contrast to our understanding of cofactor biosynthetic pathways, very little is known about cofactor catabolism. Cofactor catabolism is likely to be rare because cofactors are trace metabolites and therefore not good food sources for bacteria. Pyridoxine, **1** (vitamin B_6) catabolism is the best understood cofactor catabolic pathway and a small number of bacteria that can grow on vitamin B_6 as the sole source of carbon and nitrogen have been identified (1). Two catabolic pathways have been proposed (2). In the first pathway Figure 1, (pathway A), found in *Pseudomonas sp*. MA-1, vitamin B_6 is degraded in eight steps to form succinic semialdehyde, **9**, while in the second pathway (pathway B) observed in *Pseudomonas* IA and in *Arthrobacter* Cr-7, vitamin B_6 is catabolized in seven steps to 2-(hydroxymethyl)-4-oxobutanoate, **14**. A related catabolic pathway in *Mesorhizobium loti* MAFF303099 which is very similar to the degradative pathway A has recently been discovered. 3-hydroxy-2-methylpyridine-5-carboxylate, **7** is an intermediate in this catabolic pathway. The gene coding for the enzyme producing this metabolite has not previously been discovered and is the subject of this paper.

Previously, the genes encoding pyridoxine-4-oxidase (mlr6785), 4-pyridoxolactonase (mlr6805), pyridoxal-4-dehydrogenase (mlr6807) and the 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase (mlr6788) were identified in *Mesorhizobium loti* MAFF303099 (3-6). Recently we have reported the identification of a fifth gene (mlr6793), encoding 4-

[*]To whom correspondence should be addressed at the Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853. Telephone: (607) 255-7133. Fax: (607) 255-4137. E-mail: tpb2@cornell.edu or see3@cornell.edu
[#]These authors contributed equally to the research reported in this paper.

[‡]The Brookhaven Protein Data Bank code is 2Z7B.

pyridoxic acid dehydrogenase (7). These genes are not part of an operon, but are all close to each other on the *M. loti* chromosome, Figure 1. This suggested that other PLP catabolic genes might also be found in this region. In particular, we felt that the mlr6791gene, annotated as ribulose-5-phosphate 4-epimerase (8), was a likely candidate for the decarboxylase gene because the epimerase and the decarboxylase catalyzed reactions both proceed via enzyme stabilized enolate intermediates. In this paper, we report the cloning and overexpression of mlr6791, demonstrate that the purified gene product catalyzes the decarboxylation of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6** and describe the structure of this enzyme. While the decarboxylation of hydroxy-substituted benzene rings is a common motif in biosynthesis, the mechanism of this reaction is still poorly characterized. The structural studies described here suggest that catalysis of such decarboxylations proceeds by an aldolase-like mechanism.

EXPERIMENTAL PROCEDURES

Materials

A dehydrated form of Luria-Bertani (LB) broth was purchased from EMB Chemicals, (Gibbstown, NJ). Ampicillin and isopropyl-β-D-thiogalactopyranoside (IPTG) were obtained from Lab Scientific Inc. (Livingston, N.J.). 4-pyridoxic acid, NAD, TRIS hydrochloride, DTT and M9 minimal salts were from Sigma (St. Louis, MO.). Triethylamine was from Fisher (Fairlawn, NJ). Trifluoroacetic acid (TFA), methanol (HPLC grade), sodium chloride, imidazole, 2-mercaptoethanol and L(+)-selenomethionine were from Acros Organics (Morris Plains, NJ). Sodium dihydrogen phosphate monohydrate, calcium chloride, ferrous sulfate heptahydrate, manganese chloride and magnesium sulfate were from Mallinckrodt Baker Inc. (Phillipsburg, NJ). Deuterium oxide (D_2O) was purchased from Cambridge Isotope Laboratories Inc. (Andover, MA). Microcon YM-10 centrifugal filter devices (10,000 MWCO) and the Amicon Ultra centrifugal filter device (10,000 MWCO) were obtained from Millipore (Billerica, MA). The Supelcosil LC-18-T column for HPLC was from Supelco, (Bellefonte, PA). MEM vitamin solution, *E. coli* strain MachI and the Gateway system were from Invitrogen (Carlsbad, CA). Nucleospin Purification kit, Phusion DNA polymerase, *E. coli* BL21(DE3) and the Ni-NTA superflow resin were obtained from Macherey-Nagel (Easton, PA), New England Biolabs (Ipswich, MA), Novagen (San Diego, CA) and Qiagen (Valencia, CA) respectively.

Molecular Cloning

Standard methods were used for DNA manipulations (9), (10). Plasmid DNA was purified with the Qiagen Miniprep kit and DNA fragments were purified from agarose gel with the Nucleospin Purification kit. *Escherichia coli* strain Mach1 was used as a recipient for transformations during plasmid construction and for plasmid propagation. Phusion DNA polymerase was used for PCR following the manufacturer's recommendations. The pENTR-TEV-D-TOPO and the Gateway system were used following the manufacturer's instructions with slight modifications.

Cloning of M. loti mlr6791

The *M. loti* mlr6791 gene was amplified from genomic DNA by PCR with the following primer pair: 5′-CAC CAT GCG TCG GAA GGT CTT CGA AGA G-3′ AND 5′-TCA GGC GAG GCC TGC TTG CCT GAG G-3′. The PCR product was purified and used in a topoisomerase mediated reaction with pENTR-TEV-D-TOPO essentially following the manufacturer's instructions. Clones were screened by PCR and verified by sequencing. In initial sequencing, no clones were found with a completely correct sequence. One clone was subjected to standard site-directed mutagenesis with the following complementary primer pair: 5′-GGA TAC GTT CGG GCA CAT ATC TGC CCG TGA CCC CGA G-3′ and 5′-CTC GGG GTC ACG GGC

ACA TAT GTG CCC GAA CGT ATC C-3′. Colonies from the mutagenesis were screened with the M13-reverse primer and the mutant specific primer 5′-GAT ACG TTC GGG CAC ATA TC-3′. PCR positive clones were sequenced and a correct clone was used in an LR recombination reaction with the plasmid pDESTF1, which is a Gateway adapted vector based on the pET-system. The plasmid pDESTF1 encodes an N-terminal 6xHis tag and is under the control of the T7/lac/promoter. Clones were screened by restriction digestion. A correct clone was named pMl5335.XF1.

Overexpression and Purification

The plasmid pMl5335.XF1 was used to transform *E. coli* BL21(DE3). The cells were grown in 1L LB medium, containing 100 μg/mL of ampicillin, at 37 °C with agitation until the culture reached an OD_{590} of 0.6, at which point overexpression was induced by adding IPTG to a final concentration of 0.5 mM, the temperature was lowered to 15 °C and the cells were allowed to grow for a further 12 h before being harvested. HMPDdc was purified as described below. The yield of the purified protein was 20 mg/L.

For selenomethionine (SelMet) protein, the plasmid pMl5335.XF1 was transformed into *E. coli* B834(DE3) cells, a strain auxotrophic for methionine. Cells were grown at 37 °C with shaking in minimal M9 media which was supplemented with 20 mg/L of all amino acids except methionine, 1X MEM vitamin mix, 0.4% glucose, 50 mg/L L-selenomethionine, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 25 mg/L $FeSO_4$, and 100 mg/L ampicillin. When cells reached an $O.D._{600}$ of 0.6, the temperature was lowered to 15 °C and protein overexpression was induced with 0.5 mM IPTG. Cells were harvested after 18 h by centrifugation at 8500 *g* for 15 min. Cells were lysed by sonication and the cell lysate was cleared by centrifugation at 40000 *g* for 45 min at 4 °C. SelMet-HMPDdc was purified using a Ni-NTA affinity chromatography column. The sample was loaded in buffer A spiked with 10 mM imidazole, which contains 50 mM NaH_2PO_4, 300 mM NaCl, and 3 mM β-mercaptoethanol to prevent oxidation of the selenomethionine. The column was then washed with 10 column volumes of buffer A with 30 mM imidazole and the sample was eluted from the column with buffer A containing 250 mM imidazole. Protein was buffer exchanged into 10 mM Tris-HCl pH 7.7, 20 mM NaCl, and 1 mM DTT using an Econo-Pac 10 DG desalting column (BioRad). SelMet-HMPDdc was judged to be greater than 95% pure by SDS-PAGE gel analysis (results not shown). The yield of SelMet-HMPDdc was comparable to native protein at 20 mg per L. The protein sample was concentrated to 10 mg/mL as determined by Bradford assay using bovine serum albumin as a standard (10).

HPLC analysis

HPLC analysis of the enzymatic reaction mixture was performed on a Hewlett-Packard 1100 instrument using a Supelcosil LC-18-T (15 cm × 4.6 mm, 3.0 μM) column. Solution A contained water, solution B contained 100 mM sodium phosphate buffer at pH 6.6 and solution C contained methanol. The following linear gradient was used: 0% to 10% solution A and 100% to 90% solution B for 0 to 5 min, 10% to 48% solution A, 90% to 40% of solution B and 0% to 12% of solution C from 5-12 min, 48% to 50% solution A, 40% to 30% of solution B and 12% to 20% of solution C in 12-14 min, 50% to 30% solution A, 30% to 10% of solution B and 20% to 60% of solution C in 14-18 min, 30% to 0% solution A, 10% to 100% of solution B and 60% to 0% of solution C in 18-20 min and 0% of solution A, 100% of solution B and 0% of solution C in 20-25 minutes. The flow rate was 1 mL/min and the absorbance was measured at 254 nm (characteristic for NAD, 3-hydroxy-2-methylpyridine-5-carboxylate **7**) and 320 nm (characteristic of 4-pyridoxic acid **4**, 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6** and 3-hydroxy-2-methylpyridine-5-carboxylate,**7**). Under these conditions the following compounds were readily separated (retention time in parenthesis): 4-pyridoxic acid,

4, (17.8 min), 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6** (2.8 min), 3-hydroxy-2-methylpyridine-5-carboxylate **7 (**5.5 min) and NAD (12.9 min).

Enzymatic synthesis of the substrate

3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6** was enzymatically synthesized from 4-pyridoxic acid, **4**. The 5 mL enzymatic reaction mixture containing 12 mM NAD, 6 mM 4-pyridoxic acid, **4** and 100 μM freshly purified 4-pyridoxic acid dehydrogenase (7) in 100 mM sodium phosphate buffer at pH 8.0 was incubated overnight at room temperature. It was subsequently concentrated by lyophilizing and redissolving in a minimum volume of water. It was then filtered through YM-10 Microcon centrifugal filter at 14,000 *g* for 30 min to remove the protein and the filtrate was purified by HPLC. 100 mM ammonium acetate, pH 6.6, was used instead of 100 mM sodium phosphate buffer at pH 6.6 as solution B, to facilitate removal of the buffer salts from the isolated 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6** during lyophilization. Compound **6** is a stable white solid. ^{1}H NMR (300 MHz, D_2O) δ 2.69 (s, 3H, CH_3) and 8.42 (s, 1H, C_6-H).

Reaction time course

A time course was determined with a reaction mixture (1 mL) containing 1.58 mM 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6**, 2 μM pure enzyme and 10 μM $MnCl_2$ in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT (11). At various time points, 100 μL of the reaction mixture was quenched by addition to 100 μL of 10% TFA. This mixture was filtered through Microcon YM-10 and 100 μL of the filtrate was analyzed by HPLC, Figure 2(a).

Product purification and characterization

A reaction mixture (10.0 ml) containing 12 mM NAD, 6 mM 4-pyridoxic acid, **4**, 100 μM 4-pyridoxic acid dehydrogenase, 100 μM of HMPDdc and 1mM $MnCl_2$ in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT was incubated overnight at room temperature. It was then filtered through an YM-10 Microcon centrifugal filter at 14,000 g for 30 min. The desired enzymatic product (retention time 5.5 min) was purified by HPLC over multiple injections. Methanol was removed by rotary evaporation; TFA and water were removed under high vacuum overnight. The resulting white powder was characterized by NMR and ESI-MS. ^{1}H NMR (300 MHz, D_2O) 2.46 (s, 3H, CH_3), 7.68 (s, 1H, C4-H) and 8.12 (s, 1H, C6-H). ESI-MS (Esquire-LC_00146 instrument, Bruker, Negative ion mode) m/z = 152, (mono anionic 3-hydroxy-2-methylpyridine-5-carboxylate **7)**. Fragmentation analysis resulted in the formation of a species with m/z = 108 (M-44, decarboxylation of **7**).

Steady state kinetic parameters

The steady state kinetic parameters for HMPDdc were determined by monitoring the change in absorbance at 265 nm over time. To a reaction mixture (500 μL) containing 1 μM enzyme and 5 μM $MnCl_2$, varying concentrations of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6** were added. The rate of formation of 3-hydroxy-2-methylpyridine-5-carboxylate **7** was monitored over 3 min at 265 nm for each concentration of the substrate. The K_M and k_{cat} for the enzyme were determined by fitting the rate of product formation as a function of substrate concentration using non-linear regression to the Michaelis-Menten equation using Grafit 5.0.11 (Erithacus Software Ltd., Surray, UK). All solutions were made in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT.

Protein Crystallization

The hanging drop vapor diffusion method was used with 1 μL of SelMet protein solution and 1 μL of reservoir solution at 22 °C. Sparse matrix screens were used to determine initial hits

(Crystal Screen and Crystal Screen 2, Hampton Research). Optimized crystallization conditions consisted of 6-9% polyethylene glycol 8000 and 100 mM Tris buffer with a pH ranging from 7.0 to 7.5. Crystals grew in a conical shape to a size of approximately 0.3 mm × 0.1 mm in roughly one week and were cryoprotected by a quick transfer into crystallization conditions with an additional 17% glycerol. SelMet-HMPDdc crystals were then flash frozen by plunging in liquid nitrogen and stored frozen until data collection. Crystals belong to the space group I4 with unit cell parameters a = 72.0 Å and c = 90.4 Å. The unit cell contains one monomer per asymmetric unit with a solvent content of 45% and a Matthews coefficient of 2.24 $Å^3$/Da (12).

X-Ray Data Collection and Processing

A single wavelength anomalous diffraction dataset was collected on a SelMet-HMPDdc crystal at the NE-CAT 24-ID-C beamline at the Advanced Photon Source using a Quantum315 detector (Area Detector Systems Corp). The dataset was collected at the maximum *f*" for selenium as determined using a fluorescence scan of the SelMet-HMPDdc crystal. The crystal diffracted to 1.9 Å resolution and data was collected over 360° using a 1° oscillation range. The HKL2000 suite of programs was used to index, integrate, and scale the data (13). The data collection statistics are given in Table 1.

Structure Determination, Model Building, and Refinement

The computer program hkl2map was used to determine the positions of the Se atoms using data cut off at 2.2 Å resolution to maximize the anomalous signal (14). Three of the possible five selenium atoms were located. The program autoSHARP was used for refinement of the heavy atom positions, phasing, calculation of residual maps, density modification, and automated model building (15). Automated model building built 222 of 234 residues with the correct side chain. Examination of the density modified maps allowed the manual building of residues 139-141, 209-216, and 234 using COOT (16). Refinement of the HMPDdc model was then performed using CNS (17). A metal binding site was identified and modeled as a manganese ion, based on the magnitude of the electron density, coordination geometry, and the results of a fluorescence scan performed (results not shown). Water molecules were added as refinement continued also using CNS. The data refinement statistics are given in Table 2. The HMPDdc structure was verified using PROCHECK and no residues were located in disallowed regions of the Ramachandran plot (18). Figures were generated using Pymol (19).

RESULTS

Product purification and characterization

The product of the enzymatic reaction, was purified as a stable white powder. It was identified as 3-hydroxy-2-methylpyridine-5-carboxylate **7** by NMR and ESI-MS analysis. The presence of a singlet at 7.68 ppm in the aromatic region, corresponding to the C-4 hydrogen of **7**, indicated a clean conversion of substrate **6** to product **7**. The time course for the HMPDdc catalyzed decarboxylation of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6** to 3-hydroxy-2-methylpyridine-5-carboxylate **7**, as analyzed by HPLC is shown in Figure 2(a).

Steady state kinetic parameters

UV visible spectra of the enzymatic reaction mixture (500 μL), containing 1.58 mM of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6**, 4 μM $MnCl_2$ and 780 nM of the HMPDdc in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT, taken at various time points, Figure 2(b), showed an increase in absorbance at 265 nm. Both 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6** and 3-hydroxy-2-methylpyridine-5-carboxylate **7** absorb at this wavelength. The extinction coefficients of 3-hydroxy-2-methylpyridine-4,5-

dicarboxylate **6** and 3-hydroxy-2-methylpyridine-5-carboxylate **7** at 265 nm in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT were determined to be 1017 $M^{-1}cm^{-1}$ and 3623 $M^{-1}cm^{-1}$ respectively. The rate of product formation was determined using the difference in molar extinction coefficients (ΔA_{265} = 2615 $M^{-1}cm^{-1}$). Steady state kinetic parameters were obtained from the concentration dependence of the rate of formation of 3-hydroxy-2-methylpyridine-5-carboxylate **7** at constant concentration of HMPDdc under saturating concentration of $MnCl_2$. The enzymatic reaction exhibited Michealis Menten kinetics with K_M and k_{cat} of 366 μM and 0.6 s^{-1} respectively. The k_{cat}/K_M for HMPDdc was determined to be 1530 $M^{-1}s^{-1}$. Figure 2 (c)

Monomeric Structure of HMPDdc

The structure of HMPDdc was determined at 1.9 Å resolution using SAD phasing. All 234 residues of the protein and 3 residues from the 6X N-terminal His tag were modeled into the final structure of the monomer, as well as 205 water molecules and a manganese ion. The monomer is composed of a single domain with an αβα fold (Figure 3). The central 7 stranded mixed β-sheet is mostly antiparallel with a strand order of β5↑β6↑β7↓β4↑β1↓β2↑β3↓ where only β5 and β6 run parallel to each other as seen in Figure 3B. The β-sheet forms a half-barrel with four α-helices flanking one side of the β-sheet and two α-helices flanking the opposite side. Three of the α-helices, α1, α4 and α6, are unusually long with 19, 24 and 18 amino acids, respectively. There are also three 3_{10} helices.

Tetrameric Structure of HMPDdc

The quaternary structure of HMPDdc is a tetramer formed by using the fourfold crystallographic axis of the space group I4 and is shown in Figure 4. The tetramer is roughly 65 Å wide, 45 Å tall and 80 Å across the diagonal. A channel with a diameter of 10 Å runs through the tetramer. α-Helices α1, α4, and α6 face the channel and it is composed mostly of hydrophilic side chains and backbone carbonyl groups. The opening is much wider on the top, near the N and C termini, and nearly closed on the bottom by the last turn of α4 and the side chain of Lys194. The subunit interface is formed by interactions between the C-terminal end of one HMPDdc subunit (α4, α5, α6, and the connecting loops) and the loops between β2 and β3, β6 and β7, and the loop between the third 3_{10} helix and α3 from a neighboring subunit. The first 3_{10} helix also contributes to the interface between subunits.

Approximately 3000 $Å^2$ of surface area is buried at the interface between subunits (20). The residues lining the subunit interface are mostly hydrophobic, including a patch rich in aromatic residues, including Phe25, Phe177, Phe180, and Tyr181. There are two sets of salt bridges at the interface, one between Arg128 and Glu205 and the second between Arg162 and Glu126. The arginine residues are found close to each other and there also appears to be some stacking between these side chains. The interface is also stabilized by five hydrogen bonds: Asp23 and Thr24 both hydrogen bonding to Tyr181, Ser48 to His223, and Asp105 to Asn197. Three backbone carbonyl groups form hydrogen bonds: the carbonyl group from Ile195 to the side chain of Thr106, the carbonyl group of Asp133 to the hydroxyl group of Ser202, and the carbonyl oxygen atom of Glu134 to the oxygen of the Thr208 side chain.

Metal Binding Site

A cleft 13 Å long, 8 Å wide, and 10 Å deep forms near the hydrophobic patch at the interface between two subunits. The cleft is found in the middle of the tetramer and lies closer to the external solvent than to the channel that runs through the tetramer. This cleft contains six histidine residues in close proximity to each other: His27, His92, His94, His163, and His177 from one subunit and His113* from the neighboring subunit.

In addition to the histidine residues, a large unexpected peak in the electron density was found in this cleft and was modeled as a manganese ion. No metal was added during purification, and treatment of the protein sample with EDTA was unsuccessful in removing the metal, showing that the metal is tightly bound in *Ml*HMPDdc. A fluorescence scan identified the metal as a manganese ion. The metal has a tetrahedral coordination to His92, His94, His163, and Glu73 (Figure 5). The bonding distance between the manganese ion and His92 is 2.2 Å, while the distances to the other residues coordinated to the metal are longer than expected with a distance of 2.7 Å for His94, His163 and Glu73. Adjacent to the bound metal is a pocket 8 Å long and 7 Å wide, which could potentially accommodate 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6**. Hydrophobic residues that protrude into and line the cleft include Phe25, Phe28, and Tyr71 and hydrophilic residues in the pocket include His113*, Gln216*, and Arg219*.

DISCUSSION

Characterization of the mlr6791gene

The mlr6791 gene product was shown to catalyze the decarboxylation of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6** to 3-hydroxy-2-methylpyridine-5-carboxylate **7.** The product was purified by HPLC and characterized by NMR and ESI-MS.

Comparison of HMPDdc to Other Proteins

An iterative BLAST (21) search using the non-redundant protein sequence database revealed a large cluster of proteins with sequence identities to HMPDdc ranging from 20 to 40%. Most of these proteins are annotated as class II aldolases/adducins or hypothetical proteins. Alignment of the top hits identified five strictly conserved residues and several highly conserved residues, as shown in Supplemental Figure 1 (22,23). Pro47 is strictly conserved and is found at the interface between monomers as the first residue of a 3_{10} helix in HMPDdc. Two of the residues coordinating to the manganese ion, Glu73 and His94, are also absolutely conserved. The two other histidine residues, His92 and His163, are mostly conserved among the proteins, with His92 being replaced by an arginine and His163 being replaced by asparagine in a class II aldolase from *Nocardioides sp. JS614*. His77 is strictly conserved, and is one of the six histidine residues found near the putative active site cleft at the interface. The final conserved residue, Gly164, is adjacent to His163, which coordinates to the manganese ion and is the first residue of a β-strand.

A BLAST search was then conducted on a class II aldolase/adducin-like protein from *Burkholderia sp. 383*, which had the highest sequence similarity to *Ml*HMPDdc in the initial BLAST search. Most of the sequences with significant alignments were the same class II aldolases/adducins and hypothetical proteins identified using HMPDdc as a search subject; however, at a lower similarity level several 3,4-dihydroxyphthalate 2-decarboxylase (DHPdc) sequences were identified. DHPdc is found in the phthalate catabolism pathway of gram-positive bacteria and catalyzes a decarboxylation reaction of an aromatic ring very similar to the reaction catalyzed by HMPDdc (24,25). Alignment of these DHPdc sequences with the sequence of HMPDdc and the nearest class II aldolase identified several residues that are conserved (23). The sequence alignment prepared using ESPript (22) is shown in Supplemental Figure 2. All metal coordinating residues are conserved and His77 and His27, two histidine residues found near the cleft between subunits, are also conserved. The aromatic residues near the active site, Phe25, Tyr71, His113, and Phe138, are not conserved in DHPdc but are replaced by hydrophobic residues. Gln216 in HMPDdc is found to be an asparagine residue in DHPdc and Arg219 is replaced by a threonine residue.

The *Ml*HMPDdc monomer was submitted to DALI to identify structurally related proteins (26). Five proteins were identified as being structurally similar to HMPDdc with Z-score greater than 10; all other proteins have Z scores less than 4. Top hits included two L-fuculose 1-phosphate aldolases, one L-ribulose 5-phosphate epimerase, a rhamnulose 1-phosphate aldolase, and the C terminal fragment of phosphomethylpyrimidine kinase from *Pyrococcus furiosus* (27-29). The aldolases and epimerase belong to the superfamily of AraD-like proteins and class II aldolases. L-Fuculose 1-phosphate aldolase from *E. coli* showed the highest structural similarity with a Z score of 24.4 and an r.m.s.d. of 2.4 Å for 205 of a possible 210 residues. The sequence identity between HMPDdc and the aldolase is about 20%. The Z score for *E. coli* L-ribulose 5-phosphate epimerase was 22.5 with an r.m.s.d. of 2.4 Å for 207 of 223 residues.

The structures of the L-fuculose 1-phosphate aldolases, L-ribulose 5-phosphate epimerase, and the rhamnulose 1-phosphate aldolase superimpose well on the structure of HMPDdc, as seen in Figure 6. The topology of the long β-sheet is conserved for each of the aldolases and HMPDdc; however, HMPDdc has a C-terminal α-helix not observed in the other structures. Rhamnulose 1-phosphate aldolase is the least similar to HMPDdc and has two extra β-strands and four extra α-helices. All of the enzymes shown in Figure 6 adopt a tetrameric oligomeric state.

Despite the strong structural similarity between HMPDdc and the class II aldolase family, there is low sequence conservation and apart from the coordination of the metal ion few of the putative active site residues are conserved. The class II aldolase superfamily members each bind a zinc ion in the active site, while HMPDdc has a manganese ion bound. L-fuculose 1-phosphate aldolase from *Bacterosides thetaiotaomicron* has the highest sequence identity at 20% and the sequence identity between HMPDdc and rhamnulose 1-phosphate aldolase from *E. coli* is only 15%. The zinc ion is found in the same cleft and coordinated by three conserved histidine residues. The fourth residue coordinated to the metal ion is a glutamate residue in HMPDdc and the three aldolases, and an aspartate residue in the epimerase. Several glycine and proline residues are conserved, most likely playing roles in positioning structural elements. His77 is conserved, as is Ser29, which is hydrogen bonded to His 77. Arg31, Arg84, Pro85, and Asp86 are all structurally conserved and are found in turns directly exposed to solvent. Glu183 is conserved and is found within α4 and faces the channel that runs through the tetramer.

Active Site Comparison

The structure of L-fuculose 1-phosphate aldolase from *E. coli* with an inhibitor bound (PDB ID: 4FUA) was used to compare the active sites of the class II aldolase and HMPDdc (Figure 7) (27). The zinc ion is located closer to the coordinating residues in the aldolase structure, only 2.0 Å from His94 and His 155 and 2.1 Å from His 92. The corresponding distances in HMPDdc to the manganese ion are 2.7 Å for His 94 and His163 and 2.2 Å for His92. With the inhibitor phosphoglycolohydroxamate bound Glu73 has been pushed out of position for coordination to the zinc ion, suggesting the same could occur upon binding of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6**. His113* is replaced by Tyr113*, which points toward the zinc ion. The adjacent residue, Met114*, is found in both structures. His27 is replaced by Asn29. Phe138 is conserved, although adopting a different side chain orientation. Gly28 still composes the floor of the binding pocket and the pocket is roughly the same size in both structures. No other residues are conserved within the binding pocket, unsurprising given that the substrates are very different.

Structural Implications for Mechanism

Enzyme-catalyzed decarboxylations constitute a well-studied family of reactions and the role of pyridoxal phosphate, the pyruvoyl cofactor, imine formation with lysine and metal ions in

the catalysis of these reactions is now well established. The general rule for the catalysis of such decarboxylations is that the enzyme provides a mechanism for the stabilization of an enolate intermediate by charge delocalization (30). The decarboxylation of orotidine monophopshate that occurs during pyrimidine biosynthesis is a well-known exception (31). The decarboxylation of hydroxy-substituted benzene rings is a common motif in polyketide biosynthesis. It is generally assumed that this reaction proceeds by initial tautomerization of the phenol, followed by decarboxylation of the resulting keto-acid. However, this proposal has not been experimentally validated and only one structure of a hydroxybenzoic acid decarboxylase has been reported (PDB ID: 2DVX).

The Protein Data Bank contains three zinc-dependent decarboxylases: 2,6-dihydroxybenzoate decarboxylase (PDB ID: 2DVX), α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (PDB ID: 2HBV) and a protein of unknown function similar to α-acetolactate decarboxylase (PDB ID: 1XV2) (32); however, none of these are structurally similar to HMPDdc. In contrast several members of the class II aldolase/adducin family were revealed by a DALI search. The structural similarity between fuculose aldolase and HMPDdc suggests that the two reactions share common mechanistic features (33). For the aldolase catalyzed reaction, Tyr113 initiates the retroaldol reaction by alcohol deprotonation, Figure 8(a). The resulting enolate is stabilized by the active site zinc ion. Protonation of this enolate by Glu73 followed by product release completes the reaction. In the resting state of the enzyme, Glu73 is coordinated to the zinc ion and is released from the metal upon substrate binding. Based on this proposal, we suggest an analogous mechanism for HMPDdc. In this mechanism, binding of the substrate displaces Glu73 from the manganese ion replacing it with the substrate hydroxyl. Glu73 then provides the proton for the keto-enol tautomerization to give **23**, Figure 8(b). The decarboxylation reaction is analogous to the retroaldol reaction except that is does not need a base as the carboxylate is likely to be deprotonated under the reaction conditions. Displacement of the product from the manganese ion by Glu73 followed by product dissociation completes the reaction. The testing of this mechanistic proposal is in progress and will require additional structural and mechanistic studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENT

We thank Dr. Yasunobu Ohkawa, National Institute of Agrobiological Sciences, Ibaraki, Japan for providing us with the *M. loti*/ MAFF303090 strain; the NE-CAT beamline staff for assistance during the collection and processing of the X-ray diffraction data; Cynthia Kinsland for preparing of the overexpression plasmid and Leslie Kinsland for assisting in manuscript preparation.

This research was supported by a grant from the National Institutes of Health (GM069618).

ABBREVIATIONS

PLP, Pyridoxal 5′-phosphate
HMPDdc, 3-hydroxy-2-methyl-4,5-dicarboxylate decarboxylase
SelMet, selenomethionine
IPTG, isopropyl-β-D-thiogalactopyranoside
NAD, nicotinamide adenine dinucletide
TFA, trifluoroacetic acid
LB, Luria-Bertani
Dhpdc, 3,4-dihydroxyphthalate 2 decarboxylase
Tris HCl, Tris(hydroxymethyl)aminomethane hydrochloride

DTT, Dithiothreitol
PGH, Phosphoglycolohydroxamate

REFERENCES

(1). Rodwell VW, Volcani BE, Ikawa M, Snell EE. Bacterial oxidation of vitamin B6. I. Isopyridoxal and 5-pyridoxic acid. J Biol Chem 1958;233:1548–1554. [PubMed: 13610871]

(2). Snell, EE.; Haskell, BE. Comprehensive Biochemistry. 21. Elsevier/North Holland; New York: 1971. The Metabolism of Vitamin B6; p. 47-71.

(3). Yuan B, Yoshikane Y, Yokochi N, Ohnishi K, Yagi T. The nitrogen-fixing symbiotic bacterium Mesorhizobium loti has and expresses the gene encoding pyridoxine 4-oxidase involved in the degradation of vitamin B6. FEMS Microbiol Lett 2004;234:225–230. [PubMed: 15135526]

(4). Funami J, Yoshikane Y, Kobayashi H, Yokochi N, Yuan B, Iwasaki K, Ohnishi K, Yagi T. 4-Pyridoxolactonase from a symbiotic nitrogen-fixing bacterium Mesorhizobium loti: cloning, expression, and characterization. Biochim Biophys Acta 2005;1753:234–239. [PubMed: 16226926]

(5). Yokochi N, Nishimura S, Yoshikane Y, Ohnishi K, Yagi T. Identification of a new tetrameric pyridoxal 4-dehydrogenase as the second enzyme in the degradation pathway for pyridoxine in a nitrogen-fixing symbiotic bacterium, Mesorhizobium loti. Arch Biochem Biophys 2006;452:1–8. [PubMed: 16824480]

(6). Yuan B, Yokochi N, Yoshikane Y, Ohnishi K, Yagi T. Molecular cloning, identification and characterization of 2-methyl-3-hydroxypyridine-5-carboxylic-acid-dioxygenase-coding gene from the nitrogen-fixing symbiotic bacterium Mesorhizobium loti. J Biosci Bioeng 2006;102:504–510. [PubMed: 17270714]

(7). Mukherjee T, Kinsland C, Begley TP. PLP catabolism: Identification of the 4-Pyridoxic acid Dehydrogenase gene in Mesorhizobium loti MAFF303099. Biorganic Chemistry. 2007in Press

(8). http://theseed.uchicago.edu/FIG/protein.cgi?prot=fig|266835.1.peg.5335&user=&48hr_job=new_framework=0.

(9). Ausubel, FM.; Brent, F. Current Protocols in Molecular Biology. John Wiley and Sons; New York: 1987.

(10). Sambrook, J.; Fritsch, GF.; Maniatis, T. Molecular Cloning: A Laboratory Guide. Cold Spring Harbor Laboratory Press; Plainview, New York: 1989.

(11). Snell EE, Smucker AA, Ringelmann E, Lynen F. Bacterial Oxidation Of Vitamin B6. Iv. Enzymatic Decarboxylation Of 2-Methyl-3-Hydroxypyridine-4,5-Dicarboxylic Acid. Biochem Z 1964;341:109–119. [PubMed: 14339645]

(12). Matthews BW. Solvent content of protein crystals. J. Mol. Biol 1968;33:491–497. [PubMed: 5700707]

(13). Otwinowski Z, Minor W. Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.

(14). Pape T, Schneider TR. HKL2MAP: a graphical user interface for phasing with SHELX programs. J Appl Cryst 2004;37:843–844.

(15). Vonrhein C, Blanc E, Roversi P, Bricogne G. Automated Structure Solution With autoSHARP. Methods Mol Biol 2006;364:215–230. [PubMed: 17172768]

(16). Cowtan PE, K. Coot: Model-Building Tools for Molecular Graphics. Acta Crystallogr. D 2004;60:2126–2132. [PubMed: 15572765]

(17). Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D 1998;54:905–921. [PubMed: 9757107]

(18). Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr 1993;26:283–291.

(19). DeLano, WL. The PyMOL Molecular Graphics Systems. DeLano Scientific; San Carlos, CA: 2002.

(20). Hasel W, Hendrickson TF, Still WC. A rapid approximation to the solvent-accessible surface areas of atoms. Tetrahedron Comp. Meth 1988;1:103–16.

(21). Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–3402. [PubMed: 9254694]

(22). Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 1999;15:305–308. [PubMed: 10320398]

(23). Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680. [PubMed: 7984417]

(24). Habe H, Miyakoshi M, Chung J, Kasuga K, Yoshida T, Nojiri H, Omori T. Phthalate catabolic gene cluster is linked to the angular dioxygenase gene in Terrabacter sp. strain DBF63. Appl Microbiol Biotechnol 2003;61:44–54. [PubMed: 12658514]

(25). Eaton RW. Plasmid-encoded phthalate catabolic pathway in Arthrobacter keyseri 12B. J Bacteriol 2001;183:3689–3703. [PubMed: 11371533]

(26). Holm L, Sander C. Touring protein fold space with Dali/FSSP. Nucleic Acids Res 1998;26:316–319. [PubMed: 9399863]

(27). Dreyer MK, Schulz GE. Catalytic mechanism of the metal-dependent fuculose aldolase from Escherichia coli as derived from the structure. J Mol Biol 1996;259:458–466. [PubMed: 8676381]

(28). Luo Y, Samuel J, Mosimann SC, Lee JE, Tanner ME, Strynadka NC. The structure of L-ribulose-5-phosphate 4-epimerase: an aldolase-like platform for epimerization. Biochemistry 2001;40:14763–14771. [PubMed: 11732895]

(29). Kroemer M, Schulz GE. The structure of L-rhamnulose-1-phosphate aldolase (class II) solved by low-resolution SIR phasing and 20-fold NCS averaging. Acta Crystallogr D Biol Crystallogr 2002;58:824–832. [PubMed: 11976494]

(30). Begley TP, Ealick SE. Enzymatic reactions involving novel mechanisms of carbanion stabilization. Curr Opin Chem Biol 2004;8:508–515. [PubMed: 15450493]

(31). Miller BG, Wolfenden R. Catalytic proficiency: the unusual case of OMP decarboxylase. Annu Rev Biochem 2002;71:847–885. [PubMed: 12045113]

(32). Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242. [PubMed: 10592235]

(33). Samuel J, Luo Y, Morgan PM, Strynadka NC, Tanner ME. Catalysis and binding in L-ribulose-5-phosphate 4-epimerase: a comparison with L-fuculose-1-phosphate aldolase. Biochemistry 2001;40:14772–14780. [PubMed: 11732896]

(34). Yagi T, Kishore GM, Snell EE. The bacterial oxidation of vitamin B6. 4-Pyridoxic acid dehydrogenase: a membrane-bound enzyme from Pseudomonas MA-1. J Biol Chem 1983;258:9419–9425. [PubMed: 6348042]

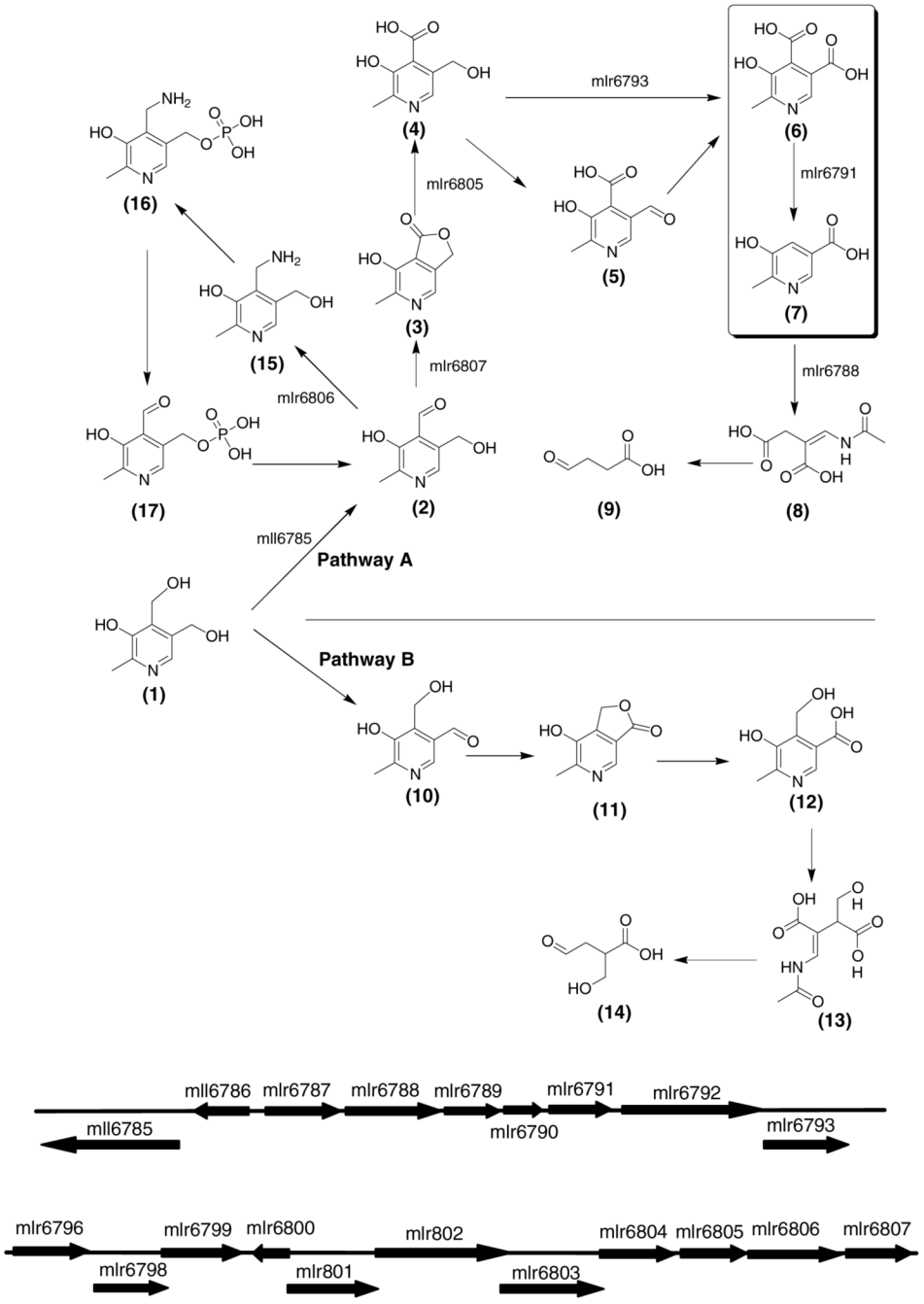

Figure 1.

The vitamin B_6 catabolic pathways. The pathway A is observed in Pseudomonas sp. MA-1 (34) while the pathway B is observed in *Pseudomonas* IA and in *Arthrobacter* Cr-7. A minor variation of pathway A is seen in *Mesorhizobium loti* MAFF303099 where the formation of **5** is not seen. The reaction shown in the box is catalyzed by HMPDdc. The bottom figure shows the gene organization in *M. loti*. The genes that have been identified to participate in PLP degradation are identified in catabolic pathway A.

1, pyridoxine; **2**, pyridoxal; **3**, 4-pyridoxolactone; **4**, 4-pyridoxic acid; **5**, 5-formyl-3-hydroxy-2-methylpyridine-4-carboxylate; **6**, 3-hydroxy-2-methylpyridine-4,5-dicarboxylate; **7**, 3-hydroxy-2-methylpyridine-5-carboxylate; **8**, 2-(acetamidomethylene) succinate; **9**,

succinic semialdehyde; **10**, isopyridoxal; **11**, 5-pyridoxolactone; **12**, 3-hydroxy-4-hydroxymethyl-2-methylpyridine-5-carboxylate; **13**, 2-(acetamidomethylene)-3-(hydroxymethyl) succinate; **14**, 2-(hydroxymethyl)-4-oxobutanoate; **15**, pyridoxamine; **16**, pyridoxamine-5-phosphate; **17** pyridoxal-5-phosphate.

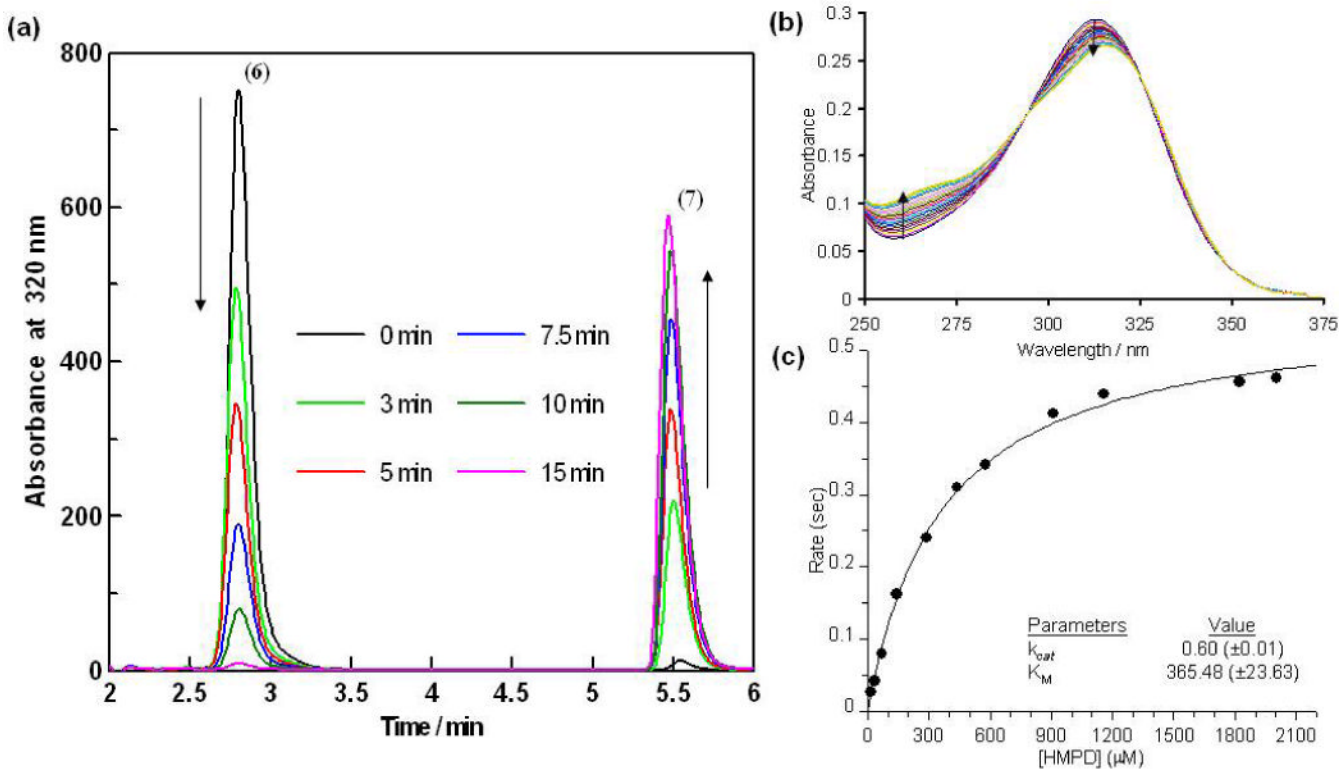

Figure 2.

(a) HPLC trace showing the disappearance of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6** (retention time of 2.8 minutes) and the appearance of 3-hydroxy-2-methylpyridine-5-carboxylate **7** (retention time of 5.5 min). (b) UV visible spectra of the enzymatic reaction mixture (500 μL) containing 1.58 mM of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6**, 4 μM $MnCl_2$ and 780 nM of the HMPDdc in 100 mM Tris HCl at pH 8.0 containing 100 mM NaCl and 2 mM DTT were taken at 1 minute intervals over 20 minutes. (c) The steady state kinetic parameters for HMPDdc were determined by monitoring the absorbance at 265 nm over time.

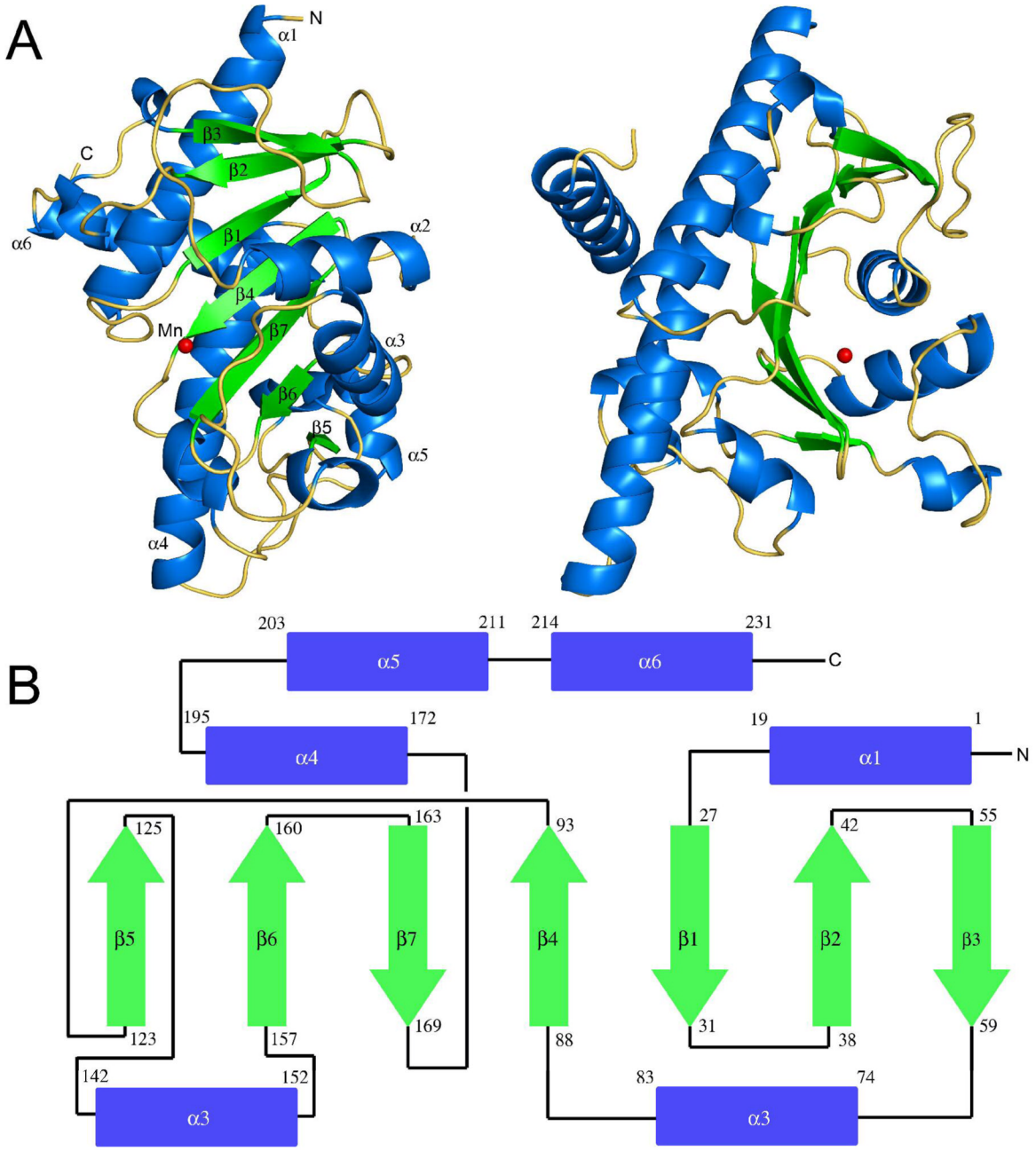

Figure 3.

Monomeric structure of *Ml*HMPDdc. A. Ribbon diagram of HMPDdc at two orientations 90° rotated from each other. The secondary structure is labeled and α-helices are colored in blue, β-strands are colored in green, and loop regions are yellow. The manganese ion is shown as a non-bonded red sphere. B. Topology diagram of HMPDdc.

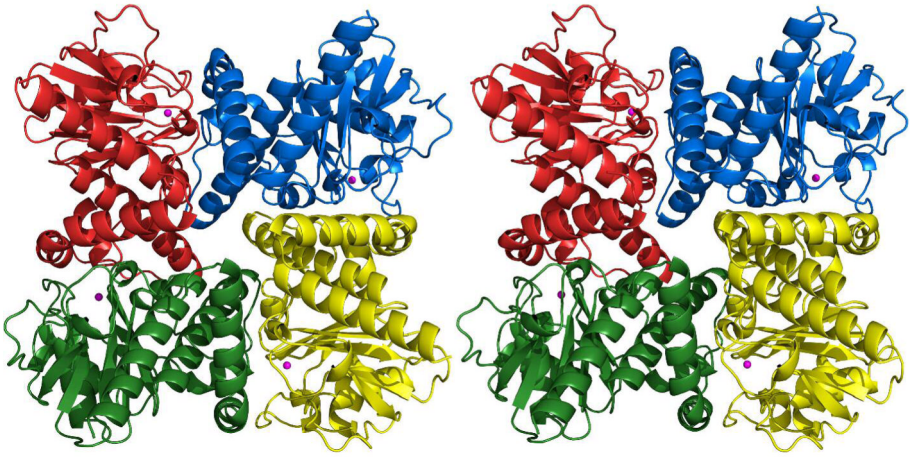

Figure 4.
Stereoview diagram of the tetrameric structure of HMPDdc. The tetramer has been color coded by subunit. The manganese ion is shown in magenta.

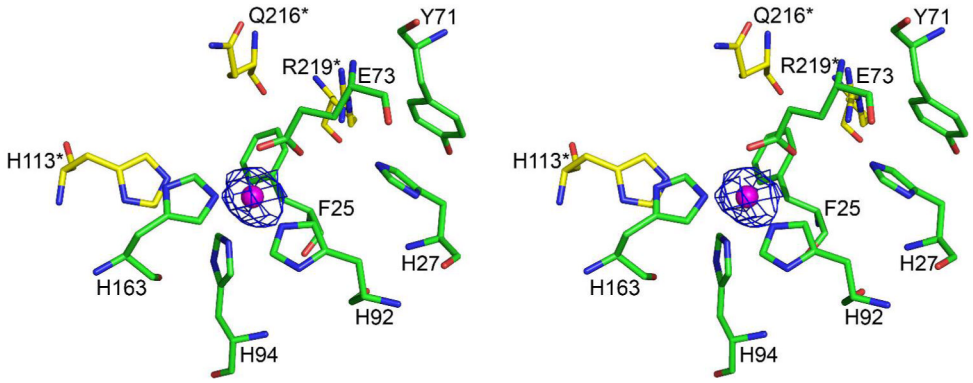

Figure 5.

Stereoview diagram of the metal binding site. Composite omit density is shown around the manganese ion at a contour level of 1.0 σ and is shown in blue. The manganese ion is magenta. Residues shown in green are from one subunit and the residues shown in yellow are from the neighboring subunit.

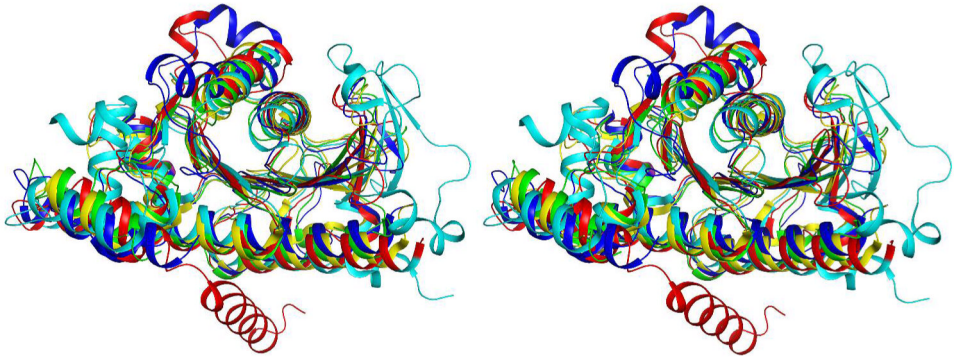

Figure 6.
Stereoview diagram of the superposition of the top four DALI hits on HMPDdc. HMPDdc is shown in red, 1-ribulose-5-phosphate-4-epimerase is shown is blue, rhamnulose-1-phosphate aldolase is in cyan, 1-fuculose-1-phosphate aldolase from *E. coli* is green, and 1-fuculose-1-phosphate aldolase from *B. thetaiotaomicron* is colored yellow.

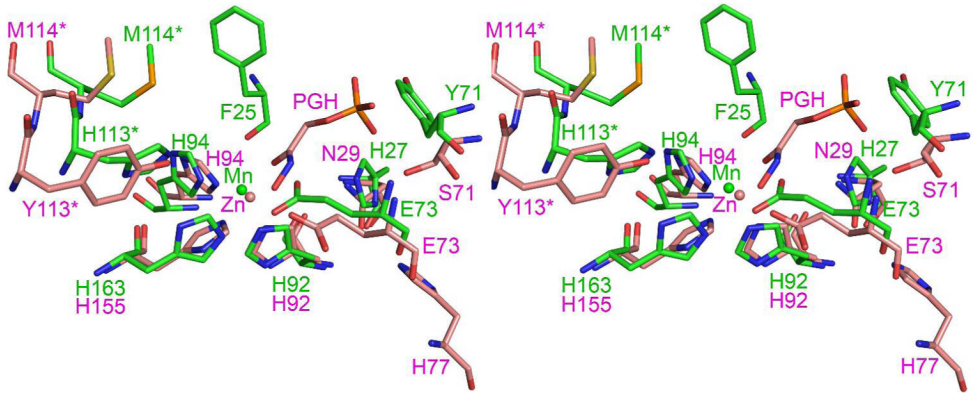

Figure 7.

Stereoview diagram of the active sites of HMPDdc and 1-fuculose1-phosphate aldolase. HMPDdc is colored in green and 1-fuculose1-phosphate aldolase is colored in pink. The phosphoglycolohydroxamate ligand is abbreviated as PGH.

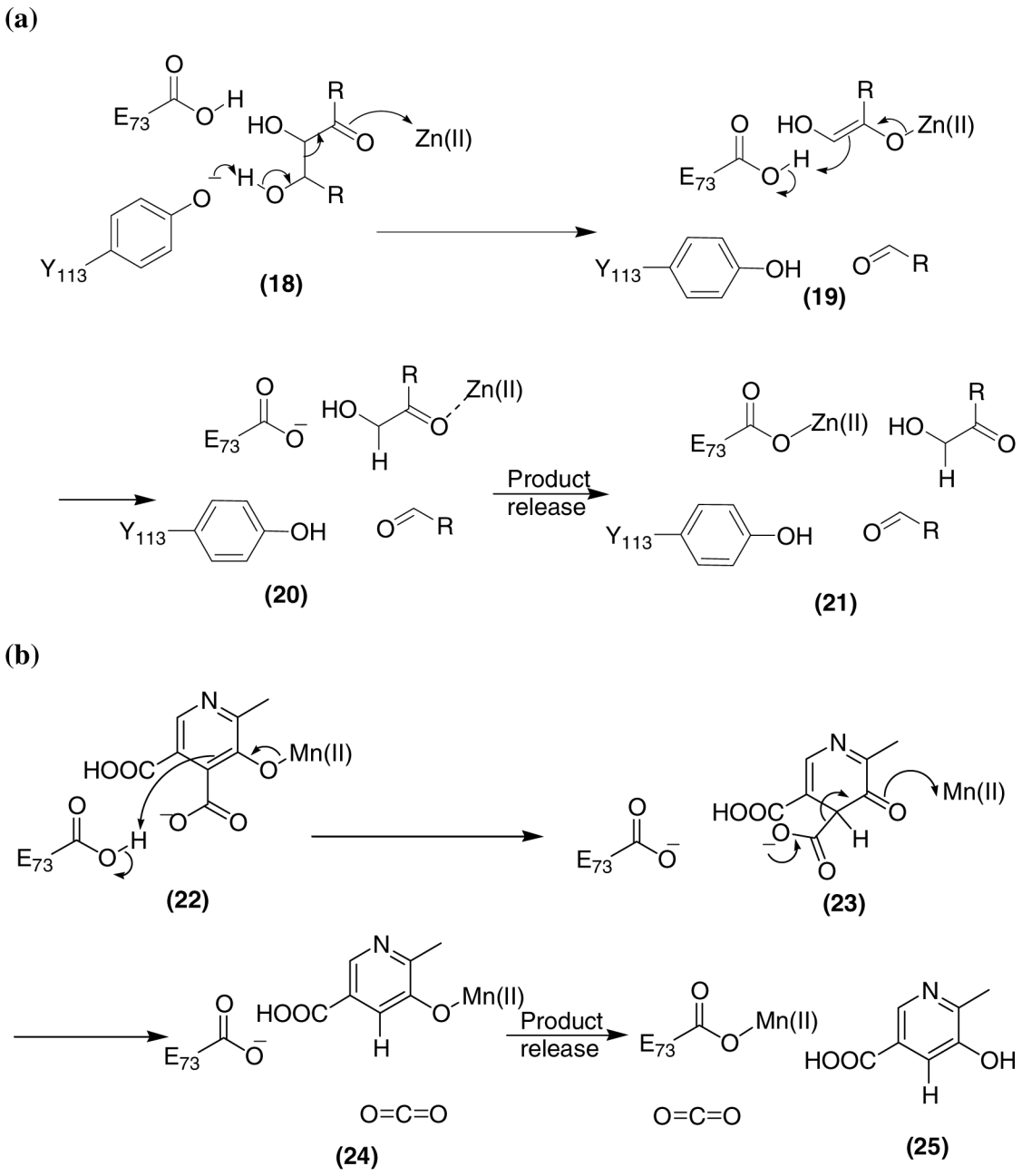

Figure 8.

(a) The retroaldol condensation reaction catalyzed by fuculose aldolase. (b) The proposed mechanism for the decarboxylation of 3-hydroxy-2-methylpyridine-4,5-dicarboxylate **6**.

Table 1

Summary of Data Collection Statistics

	SelMet HMPDdc
Beamline	APS NE-CAT 24-IDC
Resolution (Å)	1.90
Wavelength (Å)	0.97918
Space Group	*I*4
a (Å)	72.0
c (Å)	90.4
Reflections	133301
Unique reflections	34563 (2646)
Average I/σ	27.0 (4.0)
Redundancy	3.9 (2.0)
Completeness (%)	96.9 (75.1)
Rsym (%)	5.6 (14.8)

$R_{sym} = \Sigma\Sigma_i \mid I_i - \langle I \rangle \mid / \Sigma \langle I \rangle$, where <I> is the mean intensity of the N reflections with intensities I_i and common indices *h*,*k*,*l*.

Table 2

Summary of Data Refinement Statistics

	SelMet HMPDdc
Resolution (Å)	50.00 - 1.90
# of protein atoms	1852
# of metal atoms	1
# of water atoms	205
Reflections in working set	16904
Reflections in test set	906
R factor[a] (%)	19.6
Rfree[b] (%)	24.1
Rmsd from ideals	
Bonds (Å)	0.0058
Angles (°)	1.1
Avg B factor (Å^2)	30.3
Ramachandran Plot	
Most favored (%)	88.8
Additionally allowed (%)	10.7
Generously allowed (%)	0.5
Disallowed (%)	0.0

[a] R factor = Σ_{hkl} | | F_{obs} | - k | F_{cal} | | / Σ_{hkl} | F_{obs} | where F_{obs} and F_{cal} are observed and calculated structure factors, respectively.

[b] For R_{free} the sum is extended over a subset of reflections (10%) excluded from all stages of refinement.