Articles

Structural Evidence for a 1,2-Enediolate Intermediate in the Reaction Catalyzed by 3-Keto-L-Gulonate 6-Phosphate Decarboxylase, a Member of the Orotidine 5′-Monophosphate Decarboxylase Suprafamily†,‡

Eric L. Wise,[#] Wen Shan Yew,[⊥] John A. Gerlt,[⊥] and Ivan Rayment*,[#]

Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, and Departments of Biochemistry and Chemistry, University of Illinois, Urbana, Illinois 61801

Received May 27, 2003; Revised Manuscript Received August 13, 2003

ABSTRACT: 3-Keto-L-gulonate 6-phosphate decarboxylase (KGPDC) and orotidine 5′-phosphate decarboxylase (OMPDC) are members of an enzyme suprafamily, the OMPDC suprafamily, because they are homologous enzymes that catalyze mechanistically distinct reactions using different substrates. KGPDC catalyzes the Mg^{2+} ion-dependent decarboxylation of 3-keto-L-gulonate 6-phosphate to yield L-xylulose 5-phosphate and CO_2; OMPDC catalyzes the metal ion-independent decarboxylation of OMP to UMP and CO_2. Structural studies have shown that KGPDC and OMPDC share several strictly conserved active site residues that are used differently by each enzyme to catalyze their mechanistically distinct reactions. Although the mechanism of the KGPDC-catalyzed reaction has yet to be elucidated, it is thought to proceed via a Mg^{2+} ion-stabilized 1,2-enediolate intermediate. Here we report the crystal structures of KGPDC complexed with L-gulonate 6-phosphate, L-threonohydroxamate 4-phosphate, and L-xylitol 5-phosphate, analogues of the substrate, enediolate intermediate, and product, as well as with the product, L-xylulose 5-phosphate, at 1.2, 1.8, 1.7, and 1.8 Å resolution, respectively. These structures support a mechanism that involves the formation of a *cis*-1,2-enediolate intermediate. Contrary to expectations, the geometry of the intermediate does not involve bidentate coordination of both enediolate oxygen atoms to the Mg^{2+} ion but rather involves only the coordination of the oxygen on C2 to the Mg^{2+} ion. The oxygen atom on C1 instead forms hydrogen bonds to both Lys64 and Asp67, two strictly conserved active site residues. Lys64 also interacts with the oxygen on C2 and may serve to stabilize a cis conformation of the 1,2-enediolate. These structures also implicate His136 to be the general acid that protonates the 1,2-enediolate intermediate. This study further demonstrates that multiple unrelated enzyme functions can evolve from a single active site architecture without regard for substrate binding affinity or mechanism.

Understanding the strategies of natural divergent enzyme evolution requires a detailed knowledge of the relationship between structure and function in divergently related homologous enzymes. Because of the time scale involved, the evolution of new enzyme function in nature cannot be observed directly. Consequently, it is only possible to examine the results of divergent evolution and, from these, infer the mechanisms by which nature creates new enzyme functions. Careful analysis of sequence, structure, and function of divergently related enzymes has provided insight into the forces that govern divergent enzyme evolution.

Two principle strategies for the divergent evolution of enzyme function in nature are well recognized (*1*). The first involves the evolution of new enzyme function that is based on a conserved mechanistic feature or partial reaction. The second involves the evolution of new enzyme function that is based on a conserved substrate binding affinity (*2*). The predominance of these two strategies in nature underscores the close relationship between catalytic function and both chemical mechanism and substrate affinity.

Recently, a third route to the evolution of new enzymatic functions has been recognized in which neither mechanism nor substrate specificity but rather the architecture of the active site is conserved. While active-site residues are conserved, their roles are completely different (*1*, *3*). The first clear example of this evolutionary strategy was dem-

† This research was supported by Grants GM-52594 (to J.A.G. and I.R.) and GM-65155 (to J.A.G. and I.R.) from the National Institutes of Health. E.L.W. was supported by the NIH Biophysics Training Grant GM08293. Use of the Argonne National Laboratory Structural Biology Center beamline at the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Energy Research, under Contract No. W-31-109-ENG-38.

‡ The X-ray coordinates and structure factors for KGPDC have been deposited in the Protein Data Bank (1Q6O, 1Q6L, 1Q6R, and 1Q6Q for L-gulonate 6-phosphate, L-threonohydroxamate 4-phosphate, L-xylulose 5-phosphate, and L-xylitol 5-phosphate).

* To whom correspondence should be addressed. Dr. Ivan Rayment, Department of Biochemistry, 433 Babcock Drive, Madison, WI 53706. Phone: (608) 262-0437; fax, (608) 262-1319; e-mail: Ivan_Rayment@biochem.wisc.edu. Dr. John A. Gerlt, Department of Biochemistry, University of Illinois, 600 South Mathews Avenue, Urbana, IL 61801. Phone, (217) 244-7414; fax, (217) 244-6538; e-mail, j-gerlt@uiuc.edu..

[#] University of Wisconsin.

[⊥] University of Illinois.

10.1021/bi0348819 CCC: $25.00
Published on Web 10/04/2003

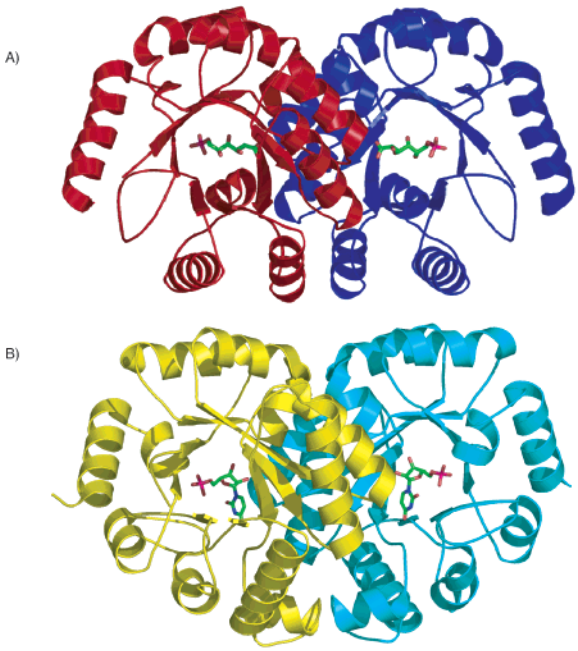

FIGURE 1: (A, B) Ribbon representations of the KGPDC and OMPDC dimers. The subunits in each protein are related to each other by a 180° axis of rotation. The locations of the two active sites are shown by the position of the L-gulonate 6-phosphate substrate analogue. As can be seen, the active site lies at the dimer interface for both enzymes. Figures 1 and 3−5 were prepared with the program Ribbons (*28*).

onstrated in the sequence and structural relationship between the 3-keto-L-gulonate 6-phosphate decarboxylase (KGPDC)[1] encoded by the UlaD gene in *Escherichia coli* and orotidine 5′-monophosphate decarboxylase (OMPDC). These enzymes, despite the slight sequence identity shared between them, lack any commonality in terms of reaction mechanism or substrate specificity (*3*, *4*). Instead, KGPDC and OMPDC share a common active-site architecture, constructed from conserved residues, that each enzyme uses in a different way to catalyze its respective reaction. Because of the lack of similarity in substrate specificity or mechanism, KGPDC and OMPDC are said to constitute an enzyme "suprafamily" (*1*).

The structures of four OMPDC's (*4*−*7*) and, more recently, KPGDC (*3*) have been reported. Given the modest pairwise sequence identities (∼20%), the degree of structural similarity between them is surprisingly high. Both enzymes adopt similar $(\beta/\alpha)_8$ folds and assemble as dimers that share a common quaternary structure (Figure 1). The active sites of OMPDC and KGPDC are remarkably alike. In both the active sites lie at the dimer interfaces and are composed from residues primarily from one subunit, but include a conserved aspartate residue (Asp67 in KGPDC, Asp 65 in OMPDC from *Bacillus subtilis*) from the symmetry related molecule in a manner that is identical for both enzymes (*3*). A constellation of conserved residues is found in nearly identical positions in the active sites of both KGPDC and OMPDC, including an aspartate at the end of the first β-strand and an Asp-X-Lys-X-X-Asp motif found at the end of the third β-strand. Considerable work has been done to understand the roles of these conserved residues in OMPDC (*8*); much less is known about these residues in KGPDC.

Despite the high degree of structural similarity between both the overall structures and active sites of KGPDC and OMPDC, there is almost no similarity in substrate specificity or reaction mechanism (Scheme 1). OMPDC accomplishes the difficult task of eliminating CO_2 from an sp^2-hybridized carbon while avoiding the formation of an energetically unfavorable vinyl anion intermediate (*4*, *7*, *9*). The mechanism of the OMPDC-catalyzed reaction is of particular interest because it is the most proficient enzyme identified to date. While the exact mechanism is still in dispute, it likely proceeds in a metal-ion independent manner in a single step, possibly via a concerted S_E2-type mechanism (a thorough discussion of the proposed mechanisms for OMPDC is given in ref *10*). KGPDC, on the other hand, requires a divalent cation for catalysis and has been proposed to catalyze the conversion of 3-keto-L-gulonate 6-phosphate to L-xylulose 5-phosphate in a stepwise manner through a 1,2-enediolate intermediate (*3*). In this mechanism, a Mg^{2+} ion has been proposed to electrostatically stabilize the negatively charged enediolate formed by the elimination of CO_2 (Scheme 1).

Scheme 1: The Overall Reactions Catalyzed by KGPDC and OMPDC

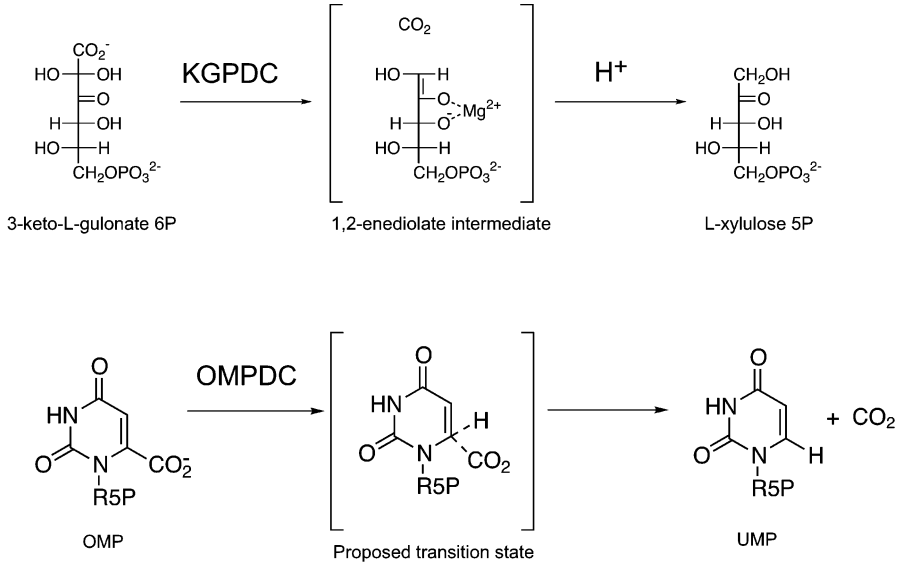

Table 1: Data Collection Statistics

	KGPDC•L-gulonate 6-P	KGPDC•L-THP	KGPDC•L-xylulose 5-P	KGPDC•L-xylitol 5-P
wavelength (Å)	0.900	1.514	0.987	0.900
source	APS 14-BC	Cu Kα	APS 19-BM	APS 14BC
unit cell dimensions	$a = 123.2$, $b = 41.8$, $c = 91.0$, $\beta = 97.2°$	$a = 123.0$, $b = 41.9$, $c = 91.5$, $\beta = 96.9°$	$a = 123.9$, $b = 41.6$, $c = 91.6$, $\beta = 96.4°$	$a = 123.2$, $b = 41.7$, $c = 92.1$, $\beta = 95.9°$
resolution (Å)[a]	1.20 (1.24−1.20)	1.8 (1.86−1.80)	1.76 (1.82−1.76)	1.69 (1.75−1.69)
unique reflections	142,337	32,215	46,548	49,731
total reflections	931,685	56,895	653,244	526,555
completeness (%)	92.4 (69.7)	91.0 (77.8)	98.2(98.1)	95.4 (82.8)
average I/σ[a]	29.6 (4.5)	5.6 (3.2)	18.1 (2.62)	27.8 (5.8)
R-merge[a,b]	0.059 (0.302)	0.030 (0.099)	0.058 (0.482)	0.06 (0.227)

[a] The value in parentheses give the statistics for the highest resolution shell which correspond to 1.24−1.20, 1.86−1.80, 1.82−1.76, and 1.75−1.69 for the complex of KGPDC with L-gulonate 6-phosphate, L-threonohydroxamate 4-phosphate, L-xylitol 5-phosphate, and L-xylulose 5-phosphate, respectively. [b] Rmerge = $(\Sigma|I_{hkl} - I|/(\Sigma I_{hkl}) \times 100$ where the average intensity I is taken over all symmetry equivalent measurements and I_{hkl} is the measured intensity for a given reflection.

Prior to the initial structural studies of KGPDC, it was assumed that the Mg^{2+} ion would coordinate the vicinal oxygens on C1 and C2 of the proposed 1,2-enediolate intermediate. The structure of KGPDC with bound L-gulonate 6-phosphate, an analogue of the substrate, however, revealed that the Mg^{2+} was coordinated to the oxygens of C3 and C4, not of C2 and C3 of the inhibitor as would be required for the expected coordination. At that time, it was unknown whether the metal ion coordination observed for the substrate analogue was the same as that of the enediolate intermediate or rather an artifact of the inhibitor. Likewise, the general acid, presumably required to deliver a solvent-derived proton to the enediolate intermediate to generate the product, had not been identified.

To understand the mechanism of the KGPDC-catalyzed reaction better, we have determined the structures of the enzyme in complex with L-gulonate 6-phosphate, an analogue of the substrate; L-threonohydroxamate 4-phosphate, an analogue of the enediolate intermediate; L-xylulose 5-phosphate, the enzymatic product; and L-xylitol 5-phosphate, an analogue of the product (Figure 2). The structure of the substrate analogue, L-gulonate 6-phosphate, described here is at a higher resolution than that previously reported (*3*). Taken together, these structures provide support for a mechanism that involves the formation of a Mg^{2+} ion-stabilized 1,2-enediolate intermediate in the KGPDC-catalyzed reaction and allow the assignment of potential catalytic roles for Lys64 and Asp67, two residues that are strictly conserved between KGPDC and OMPDC. In addition, these structures provide evidence that suggests that His136 is the general acid that protonates the enediolate intermediate. These studies highlight the dramatic mechanistic differences between OMPDC and KGPDC and provide further support for the concept of a mechanistically distinct suprafamily.

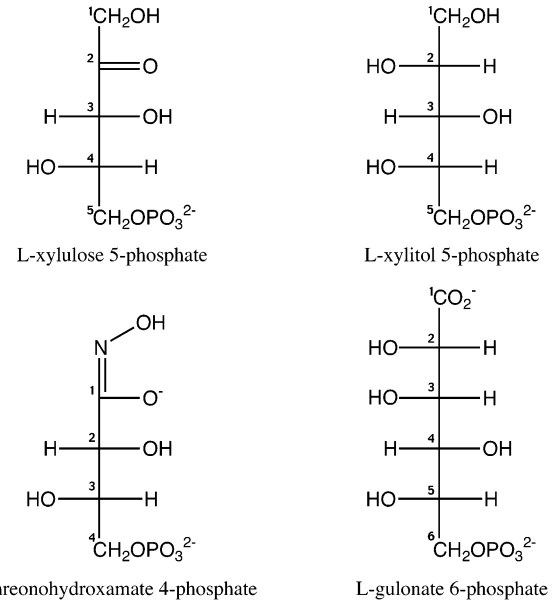

FIGURE 2: Structural formulas of the ligands examined in the present study. The numbering convention for the carbon atoms is shown.

MATERIALS AND METHODS

Protein Purification and Crystallization. KGPDC was expressed and purified as described previously (*11*). The protein was dialyzed against 50 mM HEPES, pH 7.5, 5 mM $MgCl_2$, 100 mM NaCl, concentrated to 15 mg/mL, frozen as small pellets in liquid nitrogen, and stored at −80° C. Crystals of KGPDC were grown by microbatch (*12*) by combining 10 μL of protein solution and 10 μL of a solution containing 16% MePEG 5000, 100 mM BTP pH 7.0, and 5 mM $MgCl_2$ as described before except for the presence of 25 mM inhibitor or product (*3*). All of the complexes crystallized in the space group C2 with a dimer in the asymmetric unit. Cell dimensions for each complex after freezing are given in Table 1.

Synthesis of L-*Gulonate 6-Phosphate.* L-Gulono-γ-lactone (Sigma; 5 mmol) was dissolved in 8 mL of a solution containing 50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$. A solution of NaOH (5 M) was added to increase the pH to 10−12; after 1 h the reaction was stopped by lowering the pH to 8.0 with 1 M HCl to obtain a stock solution of L-gulonate. L-Gulonate was converted enzymatically to L-gulonate 6-phosphate by adding 40 mM L-gulonate to 0.76 mM ATP, 13 mM $MgCl_2$, 50 mM potassium HEPES, pH 7.5, 31.7 mM acetyl phosphate, 450 units of L-xylulose kinase, and 270 units of acetate kinase in a volume of 250

[1] Abbreviations: PS, Advanced Photon Source at Argonne National Laboratory, Argonne, IL; KGPDC, 3-keto-L-gulonate 6-phosphate decarboxylase; OMP, orotidine 5′-monophosphate; OMPDC, orotidine 5′-monophosphate decarboxylase; HPS, D-*arabino*-hex-3-ulose 6-phosphate synthase; RMS, root-mean-square; MePEG, methyl ether poly(ethylene glycol); LTHP, L-threonohydroxamate 4-phosphate.

Table 2: Refinement Statistics

	KGPDC•L-gulonate 6-P	KGPDC•L-THP	KGPDC•L-xylulose 5P	KGPDC•L-xylitol 5-P
resolution limits (Å)	500−1.20	30−1.8	30−1.76	30−1.69
R-factor (%)	14.1	17.3	18.9	17.2
R-free (%)	16.0	21.8	22.9	20.6
no. of reflections (working set)[a]	124,961	38,709	45,980	47,245
no. of reflections (test set)[a]	6,607	2,042	2,482	2482
no. of protein atoms	3,297	3270	3,267	3,278
no. of solvent atoms	598	259	333	440
average B factor (Å^2)	14.3	24.3	26.3	20.9
RMS bond lengths (Å)	0.012	0.023	0.018	0.020
RMS bond angles (deg)	1.50	1.77	1.72	1.70

[a] The number of reflections in the test and working sets are not identical to the number of reflections listed in the data collection statistics in Table 1 because Friedel pairs were averaged during refinement.

mL. The pH of the solution was maintained at 7.5 by the addition of 1 M KOH. The reaction was allowed to proceed for 3 h at which time the enzymes were removed by ultrafiltration, and unreacted nucleotides were removed by passing the solution over activated charcoal.

Synthesis of L-*Threonohydroxamate 4-Phosphate*. Calcium L-threonate (Sigma) was converted to L-threono-1,4-lactone through passage over Dowex-50 resin (200−400 mesh, H^+-form), followed by rotary evaporation and storage in a vacuum desiccator in the presence of $Mg(ClO_4)_2$ for 4 days (*13*). An excess volume of 50% aqueous hydroxylamine was added to a solution of L-threono-1,4-lactone (1 mmol) and the solution was allowed to stir for 10 min at room temperature to yield L-threonohydroxamic acid (*14*). Unreacted hydroxylamine and solvent were removed by rotary evaporation. The product was characterized by ESI-MS as well as ^{1}H and ^{13}C NMR spectroscopy. L-Threonohydroxamate 4-phosphate was generated enzymatically by adding 20 mM L-threonohydroxamic acid to 1.32 mM ATP, 10 mM $MgCl_2$, 50 mM potassium HEPES, pH 7.5, 20 mM acetyl phosphate, 250 units of L-xylulose kinase, and 170 units of acetate kinase in a reaction mixture of 50 mL. The reaction was allowed to proceed for 2 h at which point the enzymes were removed by ultrafiltration and unreacted nucleotides were removed by passing the solution over activated charcoal.

Synthesis of L-*Xylulose 5-Phosphate*. A solution of L-xylose (15 g; Sigma) was refluxed in dry pyridine (150 mL) for 5 h after which the solvent was removed by rotary evaporation. A total of 25 mL of warm 96% ethanol was added and the solution was cooled to 4 °C overnight. Ether (25 mL) was added, and the mixture was cooled to −20 °C for 24 h. The solution was filtered to remove unreacted L-xylose, washed with cold ethanol and excess solvent was evaporated to yield a syrup of L-xylulose. L-Xylulose was converted to L-xylulose 5-phosphate enzymatically by adding 52 mM L-xylulose to 0.76 mM ATP, 13 mM $MgCl_2$, 50 mM potasium HEPES, pH 7.5, 31.6 mM acetyl phosphate, 375 units of L-xylulose kinase, and 270 units of acetate kinase in a total volume of 250 mL. The reaction was allowed to proceed for 3 h at room temperature at which time the enzymes were removed by ultrafiltration and unreacted nucleotides were removed by passing the solution over activated charcoal.

Synthesis of L-*Xylitol 5-Phosphate*. L-Xylitol was generated enzymatically by adding 40 mM xylitol (Aldrich) to 0.76 mM ATP, 13 mM $MgCl_2$, 50 mM potasium HEPES, pH 7.5, 31.6 mM acetyl phosphate, 1125 units of L-xylulose kinse, and 440 units of acetate in a volume of 250 mL. The reaction was allowed to proceed for 16 h at room temperature at which point the enzymes were removed by ultrafiltration and unreacted nucleotides were removed by passing the solution over activated charcoal.

Purification and Characterization of Ligands. Purification of the four KGPDC ligands used for crystallization experiments was accomplished by column chromatography using a DEAE Sepharose Fast Flow (HCO_3^-) column with a linear gradient of triethylammonium bicarbonate as an eluant. Fractions containing phosphate esters were identified using the Hanes-Isherwood spray reagent and TLC (*15*). Fractions containing the product were pooled and concentrated by rotary evaporation. The product in each case was characterized by ESI-MS as well as ^{1}H, ^{13}C, and ^{31}P NMR spectroscopy.

*Data Collection and Structure Determination of KGPDC•*L-*Gulonate 6-Phosphate.* Crystals grown in the presence of 25 mM L-gulonate 6-phosphate were transferred in two steps for approximately 30 s to a cryoprotecting solution containing 15% MePEG 5000, 100 mM PIPES pH 7.0, 10% ethylene glycol, 200 mM NaCl, and 50 mM L-gulonate 6-phosphate. The crystals were flash frozen in a cold nitrogen stream at −160 °C prior to data collection.

Data were collected to 1.2 Å resolution at beamline 14C of the BioCARS facility at the Advanced Photon Source, Argonne National Laboratory, Argonne IL (APS) with a single 180° scan with 1° oscillations at 20 s per frame. The data were integrated and scaled with Denzo and Scalepack (*16*). The structure was solved by molecular replacement with the program Molrep (*17*, *18*) utilizing the earlier structure of KGPDC•L-gulonate 6-phosphate as a search model (*3*). The model was improved by alternating rounds of model building and refinement with the programs TURBO and Refmac (*17*, *19*, *20*). The final R-factor was 14.1% with a final R-free of 16.0% for 5% of the data omitted from the refinement. A Ramachandran plot indicates that 95.3% of the residues fall in the most favorable regions with the remaining 4.7% falling in the additionally allowed regions as determined with the program Procheck (*21*).

*Data Collection and Structure Determination of KGPDC•*L-*Threonohydroxamate 4-Phosphate.* Crystals grown in the presence of 25 mM L-threonohydroxmate 4-phosphate were transferred to a cryoprotectant solution containing 15% MePEG 5000, 100 mM PIPES pH 7.0, 10% ethylene glycol, 200 mM NaCl, and 50 mM L-threonohydroxmate 4-phosphate in two steps. The crystals were flash frozen to −160 °C in a stream of cold nitrogen prior to data collection.

Data were collected to 1.8 Å resolution on a Rigaku Raxis IV area detector as 360 frames with 0.5° oscillations and 5 min per frame utilizing CuKα radiation from a Rigaku RU200B X-ray generator operated at 50 kV and 90 mA. The data were integrated and scaled with Denzo and Scalepack (*16*). The structure was solved by molecular replacement utilizing the structure of KGPDC·L-gulonate 6-phosphate (*3*) as a search model with the program CNS (*22*). The model was improved with alternating rounds of model building and refinement using the programs TURBO and Refmac (*19*). The final R factor was 17.3% with an R-free 21.9% for 5% of the data omitted from the structural refinement. A Ramachandran plot indicates that 94.5% of the residues fall in the most favorable regions with the remaining 5.5% falling in the additionally allowed regions as assessed by the program Procheck (*21*).

Data Collection and Structure Determination of KGPDC·L-Xylulose 5-Phosphate. Crystals grown in the presence of 25 mM L-xylulose 5-phosphate were transferred to a cryoprotectant solution containing 15% MePEG 5000, 100 mM PIPES pH 7.0, 10% ethylene glycol, 200 mM NaCl, and 50 mM L-xylulose 5-phosphate in two steps. The crystals were flash frozen in a cold nitrogen stream at −160 °C prior to data collection.

Data were collected to 1.76 Å resolution at beamline 19 μ_B of the Structural Biology Center at the APS as 180 frames with 1° oscillations at 15 s per frame. The data were integrated and scaled using the program HKL2000 (*16*). The structure was solved by molecular replacement utilizing the structure of KGPDC·L-gulonate 6-phosphate as a search model with the program Molrep in the CCP4 suite of programs. The model was improved with alternating rounds of model building and refinement with the programs TURBO and Refmac (*17*, *19*, *20*). The final R-factor was 18.9% where the final R-free was 22.9% for 5% of the data omitted from the structural refinement. A Ramachandran plot indicates that 94.3% of the residues fall in the most favorable regions, 5.5% fall in the additionally allowed regions, and 0.3% fall in the generously allowed regions as determined with the program Procheck.

Data Collection and Structure Determination of KGPDC·L-Xylitol 5-Phosphate. Crystals grown in the presence of 25 mM L-xylitol 5-phosphate were transferred to a cryoprotectant solution containing 15% MePEG 5000, 100 mM PIPES pH 7.0, 10% ethylene glycol, 200 mM NaCl, and 50 mM L-xylitol 5-phosphate in two steps. The crystals were flash frozen in a cold nitrogen stream at −160 °C prior to data collection.

Data were collected to 1.69 Å resolution at beamline 14C of the BioCARS facility at APS using a single 180° scan with 1° oscillations at 20 s per frame. The data were integrated and scaled using Denzo and Scalepack (*16*). The structure was solved by molecular replacement starting from the structure of KGPDC·L-gulonate 6-phosphate with the program CNS. The model was improved with alternating rounds of model building and refinement using the programs TURBO and Refmac. The final R-factor was 17.2% with an R-free of 20.6% for 5% of the data removed from the structural refinement. A Ramachandran plot indicates that 95.5% of the residues fall in the most favorable regions with the remaining 4.5% falling in the additionally allowed regions as determined with the program Procheck.

RESULTS

The structures of KGPDC in complex with L-gulonate 6-phosphate, L-threonohydroxamate 4-phosphate, and L-xylitol 5-phosphate, analogues of the substrate, enediolate intermediate, and product, as well as with the product, L-xylulose 5-phosphate, were determined to 1.2, 1.8, 1.7, and 1.8 Å, respectively. The structure of KGPDC in complex with L-gulonate 6-phosphate was reported earlier, but was repeated here at higher resolution to substantiate the geometry of the analogue and its coordination in the active site. Each of these structures is described separately; their collective implications for the mechanism of KGPDC-catalyzed reaction are then discussed.

Structure of KGPDC·L-Gulonate 6-Phosphate. The structure of KGPDC with L-gulonate 6-phosphate was previously reported to 2.2 Å resolution (*3*) and was determined from unfrozen crystals at 4 °C. To obtain a high-resolution structure that could be better compared to the structures described here, the structure was redetermined to 1.2 Å resolution from frozen crystals utilizing synchrotron radiation. While the overall structure and the substrate binding geometry are essentially unchanged, the higher resolution structure allows for an unambiguous assignment of hydrogen bond donors and acceptors in the active site. In particular, the nitrogen on the side chain of His136 clearly can be seen to form a hydrogen bond with the side chain of Thr114 (Figure 3a). The determination of this structure from frozen crystals also eliminates the possibility that differences between this and the other structures of KGPDC are artifacts induced by freezing the crystals.

KGPDC·L-Threonohydroxamate 4-Phosphate. The structure of KGPDC cocrystallized with L-threonohydroxamate 4-phosphate was determined to 1.8 Å resolution. The electron density for the polypeptide chains in the asymmetric unit is continuous and well-defined from Leu3 to Gly216 in both subunits. The electron density for the hydroxamate inhibitor and its associated magnesium ion is unequivocal (Figure 3b). In particular, the geometry of the hydroxamate functional group is clearly planar and thus mimics the geometry of the proposed 1,2-enediolate intermediate (Scheme 1).

The oxygen atoms analogous to those on C2 and C3 of the intermediate are observed to be coordinated to the Mg^{2+} ion instead of those analogous to the oxygen atoms on C1 and C2 as had been predicted. Instead of interacting with the Mg^{2+} ion, the oxygen on N1 of the hydroxamate forms hydrogen bonds to the polar side chains of Lys64 and Asp67. In this orientation, the hydroxamate adopts a cis conformation in which the oxygen on N1 is on the same side of the carbon−nitrogen bond as the oxygen on C1. Lys64 also interacts with the oxygen on C2 and may serve to stabilize the observed cis conformation of the hydroxamate. On the basis of this structure, we predict that the 1,2-enediolate intermediate also adopts a cis conformation as had been previously proposed. The cis conformation, however, is not the stabilized through bidentate coordination with the Mg^{2+} ion alone, but rather through the Mg^{2+} ion and the polar side chains of conserved active site residues.

When compared to the structure of KGPDC·L-gulonate 6-phosphate, the side chains of His136 and Glu112, acidic residues that could potentially deliver a solvent-derived proton to the enediolate intermediate to generate the product,

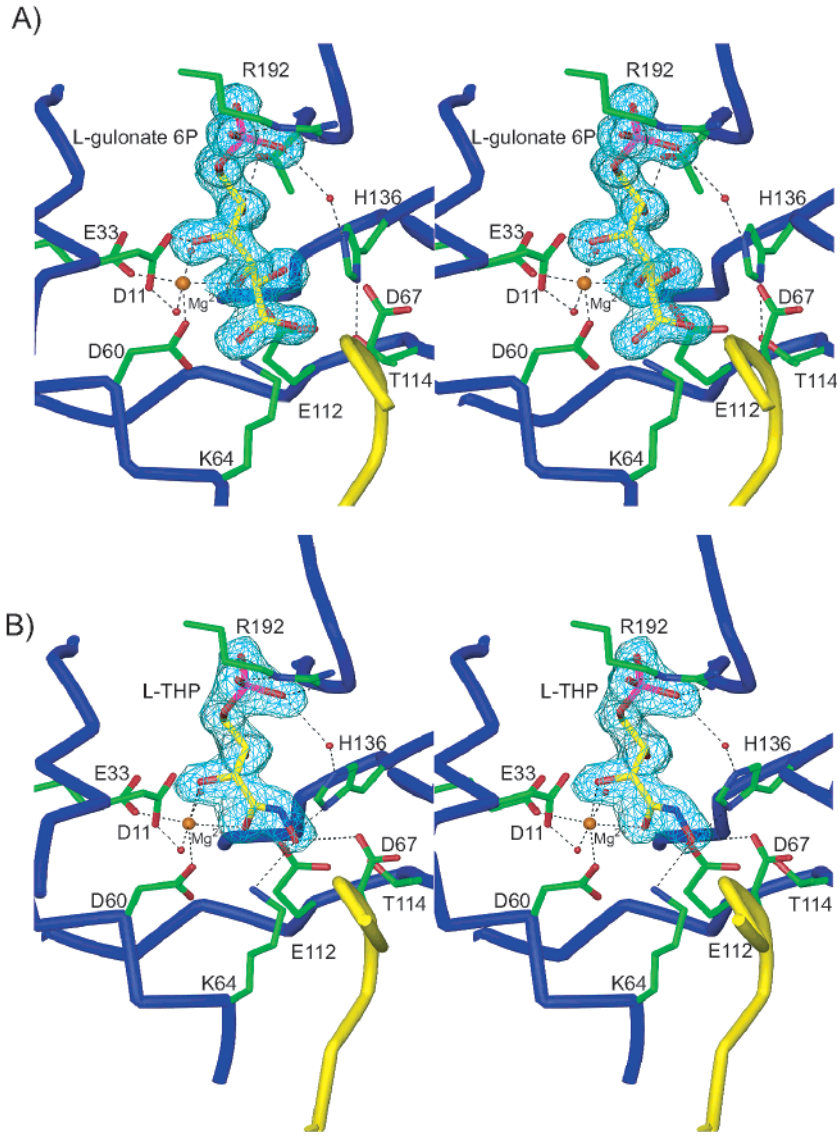

FIGURE 3: Stereoviews of the active site of KGPDC with (A) bound L-gulonate 6-phosphate, and (B) L-threonohydroxamate 4-phosphate. In both panels, the electron density for the inhibitor was depicted from a F_o-F_c map where the ligand was omitted from the phase calculation. The maps were contoured at 3 sigma. The peptide backbone for the second subunit in the active site, on which Asp67 is located, is shown in yellow. Interactions with the catalytic Mg^{2+} ion occur through the hydroxyl groups analogous to those of carbons C2 and C3 or the proposed enediolate intermediate. In the structure of the complex with L-threonohydroxamate 4-phosphate hydrogen bonds formed between the side chains of Asp67 and Lys64 function to orient and stabilize the intermediate in the active site. The side chains of His136, a potential catalytic acid, and Glu112 were observed to form a hydrogen bond. Also shown is an ordered water molecule, which forms hydrogen bonds with the His136 side chain and the phosphate group of the inhibitor, that may instead act as the catalytic acid.

have moved closer to N1 of the hydroxamate so that Nϵ of His136 is 3.5 Å from the position of C1 in the actual intermediate. Nitrogen Nϵ of the histidine side chain, in turn, forms a hydrogen bond with an ordered water molecule. This water molecule, which is present in the structure of KGPDC·L-gulonate 6-phosphate as well, interacts with an oxygen atom on the phosphate group of the inhibitor. In addition, a hydrogen bond is formed between the Nδ nitrogen of the side chain of His136 and an oxygen of the carboxylate group of Glu112. In the L-gulonate 6-phosphate structure, the side chain of His136 forms a hydrogen bond with the side chain of Thr114 rather than Glu112.

Structure of KGPDC·L-Xylulose 5-Phosphate. The structure of the complex with L-xylulose 5-phosphate, the reaction product, was determined to 1.8 Å resolution. The electron density for the protein is continuous and well-defined from Ala2 to Trp215 in both subunits. The electron density for the ligand is well defined except for oxygen on C1 as shown in Figure 4a. There is no obvious explanation for the lack of electron density for the C1 hydroxyl group. This structure, the only one described here of an actual species in the KGPDC reaction, provides further evidence that the Mg^{2+} is coordinated to the oxygens on C2 and C3 of the enediolate intermediate. As in the previously described structures, the coordination to the Mg^{2+} ion is via the oxygen atoms on C2 and C3 of the inhibitor. As in the structure of the intermediate analogue, the nitrogen on the side chain of His136 is observed to form a hydrogen bond with the side chain of Glu112.

Structure of KGPDC·L-Xylitol 5-Phosphate. The structure of KGPDC in complex with L-xylitol 5-phosphate, an analogue of the product, was determined to 1.7 Å resolution. The electron density for the protein is well-defined throughout from Leu3 to Trp215 with the exception of a stretch of

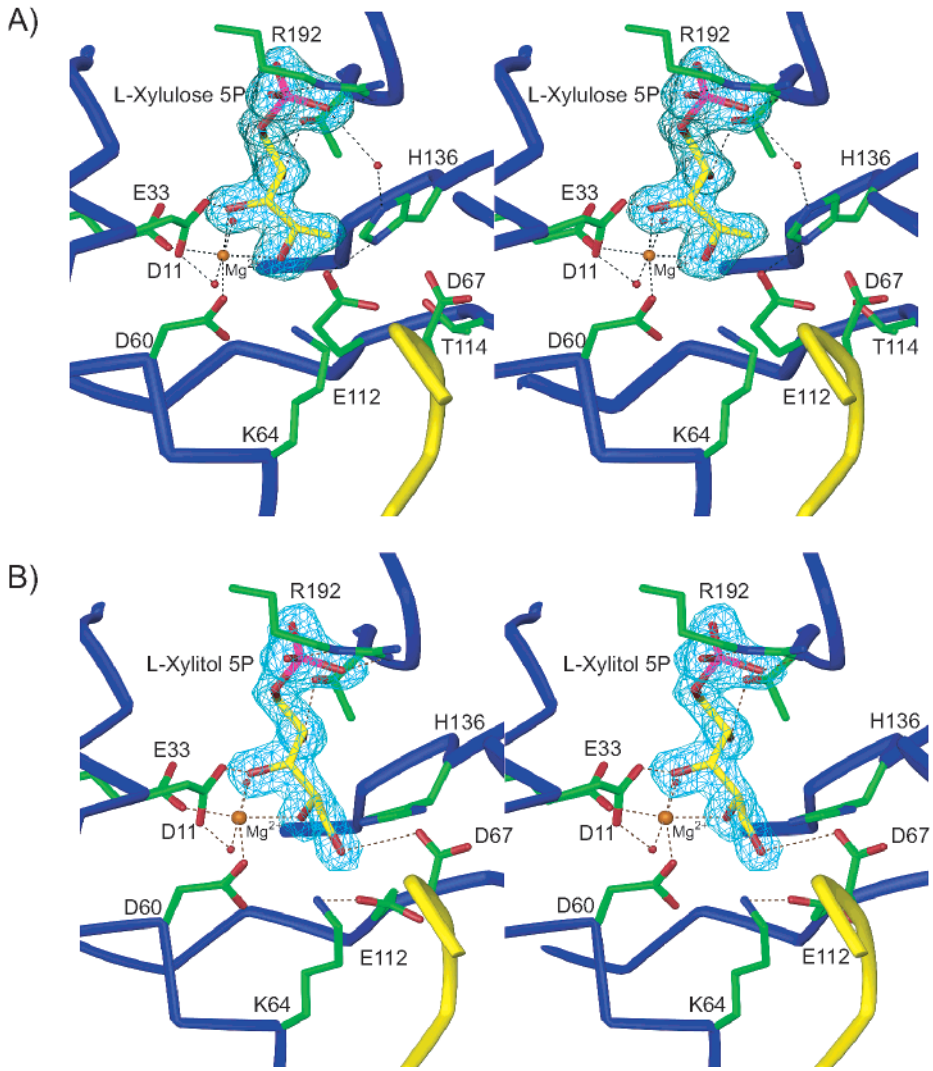

FIGURE 4: Stereoviews of the active site and electron density for (A) the product L-xylulose 5-phosphate and (B) L-xylitol 5-phosphate. In the presence of L-xylitol 5-phosphate, a product analogue, slight conformational changes occurred resulting in a rearrangement of several key active site residues. The side chains of Glu112 and His136, which form a hydrogen bond in the structures with bound intermediate and product, were observed to interact via a stacking interaction instead. The electron density for the inhibitors was depicted from maps calculated with coefficients of the form F_o-F_c map where the inhibitors were omitted from the phase calculations. The maps were contoured at 3 sigma.

residues from Ala141 to Gly145 for which only weak electron density is visible. No electron density for Gly145 in subunit A was visible and this residue not included in the structure. The coordination of the inhibitor to the Mg^{2+} ion was via the oxygen atoms on C2 and C3 of the inhibitor in a manner similar to the three previously described complexes. The electron density for the ligand is clearly defined as shown in Figure 4b. Unlike the actual reaction product, the hybridization at carbon C2 of L-xylitol 5-phosphate is sp^3 rather than sp^2. This change in hybridization prevents the ligand forming an enolate anion like the actual product and, therefore, may give additional insight into the mechanism of protonation of the enediolate intermediate.

Unlike the structure of KGPDC with the reaction product, L-xylulose 5-phosphate, the hydroxyl group of carbon C1 can clearly be seen and is observed to form a hydrogen bond with the side chain of Asp67. In addition, the side chains of Glu112 and His136 both move substantially so that instead of forming an interresidue hydrogen bond, they are stacked upon one another. These movements accompany a rearrangement of the loop that follows Glu112 and a slight shift of the entire helix that follows His136. This helix, as a whole, is less ordered than had been observed in previous structures. The movement of Glu112 and the loop which follows it causes the side chain of Thr114 to turn away from the active site (Figure 5).

DISCUSSION

Evidence for a 1,2-Enediolate Intermediate. The requirement of a divalent metal ion for catalysis strongly suggests that the likely mechanism of KGPDC involves the formation of an enediolate intermediate, in which the negative charge of the intermediate is stabilized by a Mg^{2+} ion in the active site. Despite the evidence for a metal-ion assisted mechanism, a Schiff-base mechanism, similar to that in the acetoacetate decarboxylase-catalyzed reaction, could not be excluded a priori because of the presence of the conserved Lys64 in the active site of KGPDC (*23*). Here, the structure of KGPDC with a bound analogue of the proposed intermediate provides persuasive evidence that this is not the case and that the reaction does proceed via an enediolate intermediate. In

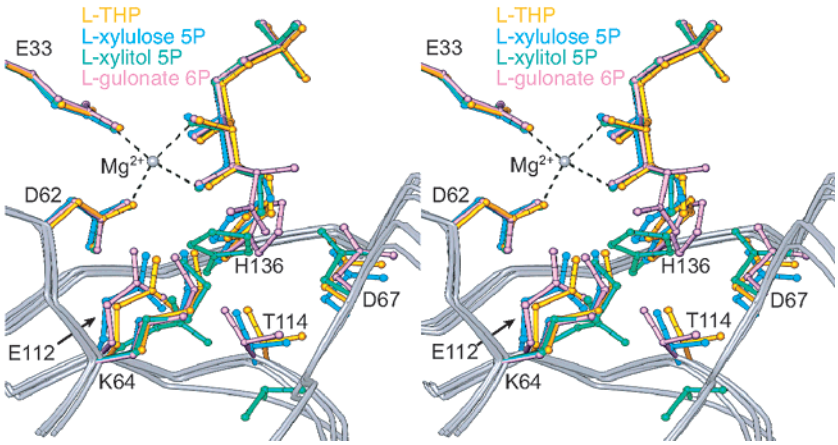

FIGURE 5: Overlay of key active site residues for the structures of KGPDC with L-gulonate 6-phosphate (lavender), L-threonohydroxamate 4-phosphate (yellow), L-xylitol 5-phosphate (green), and L-xylulose 5-phosphate (teal). Structures were aligned using the program LSQKAB in the CCP4 package (*17*).

addition, the structure of KGPDC with the product, L-xylulose 5-phosphate, showed no evidence of Schiff base formation as might be expected if that were indeed the reaction mechanism.

Stabilization of the Enediolate Intermediate. It was initially proposed that the coordination of the metal ion would be though the oxygen atoms on both carbons C1 and C2 so that the enediolate would be locked in a cis conformation. In the KGPDC·L-gulonate 6-phosphate structure, however, the oxygen atoms on C3 and C4 of the substrate analogue were observed to coordinate the metal ion instead of the oxygen atoms on C2 and C3 as had been predicted. At that time, it was believed that this anomaly may have been an artifact from the altered structure of the substrate analogue when compared to the actual substrate.

From the structures presented here, however, it is clear that the coordination of the metal ion observed in the L-gulonate 6-phosphate structure is genuine and does not differ for the reaction intermediate or product. While it would seem logical to assume that the metal ion would bridge the oxygen atoms of carbons C1 and C2 of the intermediate in a manner similar to that seen in the structures of class II aldolases (*24*), such an arrangement is not required since the negative charge generated by tautomerization of the keto group is localized primarily to the oxygen atom located on carbon C2. Consistent with earlier predictions, the enediolate appears to exist in a cis conformation. Contrary to earlier predictions, however, instead of interactions with the Mg^{2+} ion, hydrogen bonds formed between the side chain of Asp67 and the oxygen atom of carbon C1 of the intermediate and the side chain of Lys64 and the oxygen atom of carbon C2 of the intermediate serve to maintain the cis conformation. These interactions may also serve to orient the intermediate so that the enediolate is properly positioned for the addition of a proton to yield L-xylulose 5-phosphate. The binding of the C3 hydroxyl group of the intermediate to the metal ion, while not directly involved in catalysis, might function to orient the enediolate and assist in the positioning of the substrate in the active site.

Identity of a General Acid. The structure of KGPDC with a bound analogue of the enediolate intermediate allows a confident prediction to be made regarding the identity of the general acid that delivers a solvent-derived proton to the intermediate to generate the L-xylulose 5-phosphate product. The imidazole group of His136 is the only residue in the active site appropriately located to deliver a proton to the intermediate. Neither Asp67 nor Lys64 are sufficiently close to carbon C2 to participate in proton delivery and both residues form hydrogen bonds with O1 of the intermediate, which would make it unlikely that they could function as proton donors. Likewise, neither of the water molecules coordinated by the Mg^{2+} ion is positioned correctly in the active site to function as the proton donor. The position of His136 suggests that it likely acts as the general acid to protonate the intermediate from the *si* face. The Nϵ nitrogen atom of the imidazole ring is 3.5 Å from the *si* face of the double bond bridging C1 and C2 of the intermediate so that a proton on this nitrogen could easily deliver a proton to C1. Alternatively, it is also possible that the proton might be transferred to the *si* face from His136 via a closely associated water molecule. As noted earlier this well-defined solvent molecule is coordinated by both His136 and substrate phosphate moiety. With either of these scenarios, we predict that the carboxylate group of the substrate departs from the *re* face of the intermediate, the protonation occurs on the *si* face, and the overall decarboxylation reaction proceeds with an inversion of configuration.

The hydrogen bond between Nδ of His136 and the side chain of Glu112 suggests that Glu112 may act to orient the imidazole ring for addition of a proton to the intermediate. This interaction may also serve to stabilize the protonated state for His136. In the structure of KGPDC with a substrate analogue, L-gulonate 6-phosphate, the side chain of His136 was observed to form a hydrogen bond with the side chain of Thr114. This change in the state of hydrogen bonds in the active site, which accompanies the transition from the substrate to the intermediate, likely allows for the movement of the His136 side chain to a location that is optimal for addition to the intermediate (Figure 5). As would be expected, the locations of the side chains of His136 and Glu112 are closer to the position of carbon C2 of the hydroxamate than in the structures of either the substrate analogue or the product, which suggests that these two residues may move in concert to act as the proton donor to the intermediate (Figure 3).

Further evidence for such movement comes from the structure of KGPDC with a product analogue, L-xylitol 5-phosphate. In this structure, the side chains of Glu112 and

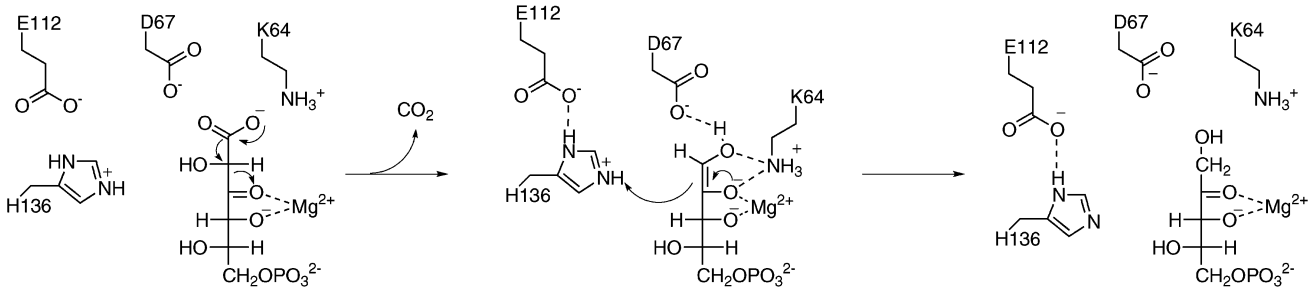

FIGURE 6: Proposed mechanism for KGPDC. Stabilization of the *cis*-1,2-enediolate intermediate provides the driving force for the decarboxylation reaction. The Mg^{2+} ion stabilizes the anionic intermediate. The side chains of two conserved residues, Lys64 and Asp67, act to orient and stabilize the intermediate. The side chain of His136, in conjunction with that of Glu112, acts as the catalytic acid to protonate the intermediate and generate the product, L-xylulose 5-phosphate.

His136 have both moved substantially so as to change the interactions of these two residues in the active site. Instead of forming a hydrogen bond as in previous structures, Glu112 and His136 are stacked on top of each other in the L-xylitol 5-phosphate structure. In this structure and in all other KGPDC structures, these regions have been more disordered and have had higher B-factors than the rest of the protein, which further underscores the mobility of these residues. While the significance of this movement is not entirely clear, it may be that this structure is analogous to an altered protonation state of His136 that occurs along the reaction pathway. In any event, it is clear that His136 has considerable conformational variability in the active site depending on the nature of the ligand, yet when confronted with an enediolate intermediate, adopts a position where it may be able to participate in proton transfer.

Proposed Mechanism for KGPDC. A mechanism for the decarboxylation of 3-keto-L-gulonate 6-phosphate by KGPDC can be envisioned to first involve the loss of the C1 carboxylate group as CO_2 to form a *cis*-1,2-enediolate intermediate (Figure 6). The driving force behind the reaction is presumably the stabilization of the intermediate by the essential Mg^{2+} ion and the conserved side chains of Lys64 and Asp67 in the active site. In addition, a hydrophobic pocket in the active site, formed by Thr36, Ile37, Ala68′, and Leu72′ (where Ala68′ and Leu72′ are contributed by the other subunit of the dimer) may destabilize the ground state by surrounding the negatively charged carboxylate group of the substrate.

Next, the imidazolium group of His136 (or a water molecule associated with His136) delivers a solvent-derived proton to the enediolate intermediate to yield the product, L-xylulose 5-phosphate. The side chain of Glu112 may function to orient the imidazole ring. Clearly, however, the details of this step of the reaction have not been clearly defined and additional work will be necessary to fully understand the interplay between these two residues.

Implications for Divergent Enzyme Evolution. While clearly homologous, KGPDC and OMPDC do not, unlike previous examples of homologous enzymes, share either a conserved substrate specificity or mechanistic feature (*25*, *26*). Both enzymes, however, share a common active site architecture that is constructed from several strictly conserved residues. KGPDC makes use of these active site residues to catalyze the decarboxylation of 3-keto-L-gulonate 6-phosphate in a metal ion-dependent manner while OMPDC uses many of the same residues to catalyze the decarboxylation of OMP via an unrelated metal ion-independent mechanism. Currently available structure−function relationships for both OMPDC and KGPDC indicate that the individual functions of these conserved active-site residues vary dramatically from one enzyme to the other. Previous studies have shown, for example, that the residues involved in the coordination of the Mg^{2+} ion in the KGPDC active site have completely different roles in the OMPDC reaction (*3*, *27*).

The structures presented here provide further evidence for the varied roles of these conserved residues in the KGPDC- and OMPDC-catalyzed reactions. In OMPDC, the side chain of Lys62 almost certainly acts as the proton donor; in KGPDC, the homologous residue, Lys64, is involved in the stabilization of the anion intermediate, whereas His136 is proposed to act as the proton donor. In OMPDC, the side chain of Asp65 is not directly involved in catalysis but rather assists in the binding of the sugar moiety of OMP(*8*); in KGPDC, the homologous residue, Asp67, is directly involved in catalysis through the stabilization and orientation of the intermediate. The importance of this residue, the second Asp in the conserved Asp-X-Lys-X-X-Asp motif at the end of the third β-strand, to both KGPDC and OMPDC cannot be understated since, in both enzymes it is the only conserved active site residue that is donated by the second subunit of the dimer to the active site. The importance of this residue to both reactions confirms the evolutionary relationship between KGPDC and OMPDC. While the function of this residue varies dramatically from OMPDC to KGPDC, its importance in both reactions may be the driving force behind the conservation of the arrangement of the dimer and the overall active site architecture in KGPDC on OMPDC.

CONCLUSIONS

The structures described in this paper provide strong evidence that the KGPDC-catalyzed decarboxylation of 3-keto-L-gulonate 6-phosphate proceeds via an *cis*-1,2-enediolate intermediate. In addition to the essential Mg^{2+} ion, two conserved active site residues, Lys64 and Asp67, act to orient and stabilize this intermediate. Although the structural data are insufficient to conclusively identify the general acid that protonates the enediolate intermediate, they point to the involvement of His136 (either directly or indirectly via an intervening water molecule). KGPDC shares a common active site architecture with the homologous OMPDC, but each enzyme uses conserved active site residues in different ways to catalyze unrelated reactions on unrelated substrates.

Clearly, nature is opportunistic when it comes to the divergent evolution of new enzyme activity: catalysts for new reactions can evolve without retaining any mechanistic characteristics or ligand specificity from the progenitor enzyme. Even though a conserved active site architecture is shared by both KGPDC and OMPDC, each enzyme uses it in different ways to catalyze different reactions.

REFERENCES

1. Gerlt, J. A., and Babbitt, P. C. (2001) *Annu. Rev. Biochem. 70*, 209−246.
2. Hocker, B., Beismann-Driemeyer, S., Hettwer, S., Lustig, A., and Sterner, R. (2001) *Nat. Struct. Biol. 8*, 32−6.
3. Wise, E., Yew, W. S., Babbitt, P. C., Gerlt, J. A., and Rayment, I. (2002) *Biochemistry 41*, 3861−9.
4. Appleby, T. C., Kinsland, C., Begley, T. P., and Ealick, S. E. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 2005−10.
5. Harris, P., Navarro Poulsen, J. C., Jensen, K. F., and Larsen, S. (2000) *Biochemistry 39*, 4217−24.
6. Miller, B. G., Hassell, A. M., Wolfenden, R., Milburn, M. V., and Short, S. A. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 2011−6.
7. Wu, N., Mo, Y., Gao, J., and Pai, E. F. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 2017−22.
8. Miller, B. G., Snider, M. J., Wolfenden, R., and Short, S. A. (2001) *J. Biol. Chem. 276*, 15174−6.
9. Sievers, A., and Wolfenden, R. (2002) *J. Am. Chem. Soc. 124*, 13986−7.
10. Miller, B. G., and Wolfenden, R. (2002) *Annu. Rev. Biochem. 71*, 847−85.
11. Yew, W. S., and Gerlt, J. A. (2002) *J. Bacteriol. 184*, 302−6.
12. Rayment, I. (2002) *Structure (Camb) 10*, 147−51.
13. Angelotti, T., Krisko, M., Oconnor, T., and Serianni, A. S. (1987) *J. Am. Chem. Soc. 109*, 4464−4472.
14. Salmon, L., Prost, E., Merienne, C., Hardre, R., and Morgant, G. (2001) *Carbohydr. Res. 335*, 195−204.
15. Hanes, C. S., and Isherwood, F. A. (1949) *Nature 164*, 1107−1112.
16. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
17. CCP4. (1994) *Acta Crystallogr. D50*, 760−763.
18. Vagin, A., and Teplyakov, A. (2000) *Acta Crystallogr. D Biol. Crystallogr. 56*, 1622−4.
19. Roussel, A., and Cambillau, C. (1991) in *Silicon Graphics Geometry Partners Directory*, Silicon Graphics, Mountain View, California.
20. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr. 53*, 240−255.
21. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
22. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−21.
23. Highbarger, L. A., Gerlt, J. A., and Kenyon, G. L. (1996) *Biochemistry 35*, 41−46.
24. Hall, D. R., Leonard, G. A., Reed, C. D., Watt, C. I., Berry, A., and Hunter, W. N. (1999) *J. Mol. Biol. 287*, 383−94.
25. Todd, A. E., Orengo, C. A., and Thornton, J. M. (2001) *J. Mol. Biol. 307*, 1113−43.
26. Copley, R. R., and Bork, P. (2000) *J. Mol. Biol. 303*, 627−41.
27. Wu, N., Gillon, W., and Pai, E. F. (2002) *Biochemistry 41*, 4002−11.
28. Carson, M. (1997) *Macromol. Crystallogr. Pt B 277*, 493−505.

BI0348819