Articles

Evolution of Enzymatic Activities in the Enolase Superfamily: D-Mannonate Dehydratase from *Novosphingobium aromaticivorans*[†,‡]

John F. Rakus,[§] Alexander A. Fedorov,[⊥] Elena V. Fedorov,[⊥] Margaret E. Glasner,[#] Jacob E. Vick,[§] Patricia C. Babbitt,[#] Steven C. Almo,*[,⊥] and John A. Gerlt*[,§]

Departments of Biochemistry and Chemistry, University of Illinois at Urbana-Champaign, 600 S. Mathews Avenue, Urbana, Illinois 61801, Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, and and Department of Biopharmaceutical Sciences, School of Pharmacy and California Institute for Quantitative Biomedical Research, University of California, 1700 4th Street, San Francisco, California 94158

Received August 21, 2007; Revised Manuscript Received September 6, 2007

ABSTRACT: The D-mannonate dehydratase (ManD) function was assigned to a group of orthologous proteins in the mechanistically diverse enolase superfamily by screening a library of acid sugars. Structures of the wild type ManD from *Novosphingobium aromaticivorans* were determined at pH 7.5 in the presence of Mg^{2+} and also in the presence of Mg^{2+} and the 2-keto-3-keto-D-gluconate dehydration product; the structure of the catalytically active K271E mutant was determined at pH 5.5 in the presence of the D-mannonate substrate. As previously observed in the structures of other members of the enolase superfamily, ManD contains two domains, an N-terminal $\alpha+\beta$ capping domain and a $(\beta/\alpha)_7\beta$-barrel domain. The barrel domain contains the ligands for the essential Mg^{2+}, Asp 210, Glu 236, and Glu 262, at the ends of the third, fourth, and fifth β-strands of the barrel domain, respectively. However, the barrel domain lacks both the Lys acid/base catalyst at the end of the second β-strand and the His-Asp dyad acid/base catalyst at the ends of the seventh and sixth β-strands, respectively, that are found in many members of the superfamily. Instead, a hydrogen-bonded dyad of Tyr 159 in a loop following the second β-strand and Arg 147 at the end of the second β-strand are positioned to initiate the reaction by abstraction of the 2-proton. Both Tyr 159 and His 212, at the end of the third β-strand, are positioned to facilitate both syn-dehydration and ketonization of the resulting enol intermediate to yield the 2-keto-3-keto-D-gluconate product with the observed retention of configuration. The identities and locations of these acid/base catalysts as well as of cationic amino acid residues that stabilize the enolate anion intermediate define a new structural strategy for catalysis (subgroup) in the mechanistically diverse enolase superfamily. With these differences, we provide additional evidence that the ligands for the essential Mg^{2+} are the only conserved residues in the enolase superfamily, establishing the primary functional importance of the Mg^{2+}-assisted strategy for stabilizing the enolate anion intermediate.

The members of the mechanistically diverse enolase superfamily catalyze reactions that are initiated by base-catalyzed abstraction of the α-proton of a carboxylate anion substrate to generate an enediolate intermediate that is stabilized by coordination to an essential Mg^{2+} (*5*−*8*). The active sites are at the interface between an N-terminal $\alpha+\beta$ capping domain and a C-terminal $(\beta/\alpha)_7\beta$-barrel domain [modified $(\beta/\alpha)_8$- or TIM[1]-barrel]. In those members that have been structurally and functionally characterized, the acid/base catalysts and ligands for the Mg^{2+} are at the C-termini of the β-strands in the barrel domain; most of the residues that determine substrate specificity are in the capping domain. The location of the catalytic groups at the ends of the β-strands in the barrel domain enables the evolution of "new" functions (*9*): the pseudosymmetric structure provides

[†] This research was supported by GM-52594, Program Project GM-71790, and the Ruth L. Kirschstein National Research Service Award 5 T32 GM070421 from the National Institutes of Health.

[‡] The X-ray coordinates and structure factors for SeMet wild type ManD liganded with Mg^{2+}, the K271E mutant of ManD liganded with Mg^{2+} and D-mannonate, and wild type ManD liganded with Mg^{2+} and 2-keto-3-deoxy-D-gluconate have been deposited in the Protein Data Bank (PDB accession codes, 2QJJ, 2QJM, and 2QJN, respectively).

* To whom correspondence should be addressed: J.A.G.: Phone: (217) 244-7414. Fax: (217) 244-6538. E-mail: j-gerlt@uiuc.edu. S.C.A.: Phone: (718) 430-2746. Fax: (718) 430-8565. E-mail: almo@aecom.yu.edu.

[§] University of Illinois at Urbana-Champaign.

[⊥] Albert Einstein College of Medicine.

[#] University of California.

[1] Abbreviations: FucD, L-fuconate dehydratase; GalD, D-galactonate dehydratase; GlucD, D-glucarate/L-idarate dehydratase; GlucDRP, GlucD-related protein; KdgK, 2-keto-3-deoxy-D-gluconate kinase; LB, Luria−Bertani broth; LDH, L-lactate dehydrogenase; MAL, β-methyl aspartate ammonia lyase; ManD, D-mannonate dehydratase; MLE, muconate lactonizing enzyme; MR, mandelate racemase; PK, pyruvate kinase; TarD, D-tartrate dehydratase; TalrD/GalrD, L-talarate/galactarate dehydratase; TIM, triose phosphate isomerase.

10.1021/bi701703w CCC: $37.00
Published on Web 10/18/2007

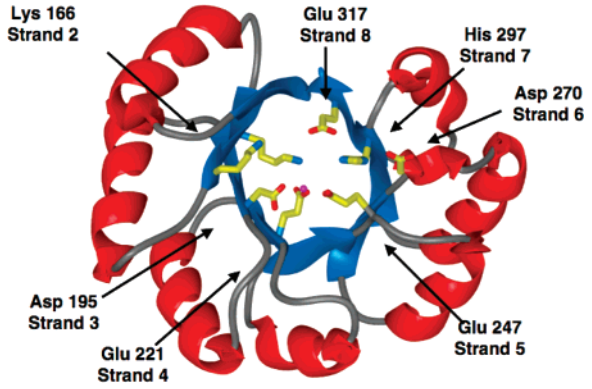


FIGURE 1: The $(\beta/\alpha)_7\beta$-barrel domain of mandelate racemase (MR) showing the conserved positions of acid/base catalysts and ligands for the Mg^{2+} at the C-termini of the β-strands.

a versatile scaffold so that functional groups can be positioned on either face of the enediolate intermediate, thereby allowing catalysis of reactions with differing stereochemical requirements. Figure 1 illustrates the conserved positions of the acid/base catalysts and ligands for the Mg^{2+} at the ends of the β-strands in the $(\beta/\alpha)_7\beta$-barrel domain using mandelate racemase as an example.

Five structural strategies for catalysis (subgroups) have been identified by structural and bioinformatic analyses of functionally characterized members,[2] all of which share positionally conserved ligands for the essential Mg^{2+} at the ends of the third, fourth, and fifth β-strands in the barrel domain (Table 1 contains a summary of the conserved resides in the barrel domains in each subgroup):

(1) In the enolase subgroup (~1200 sequences[3]), the Lys that initiates the reaction by abstraction of the 2-proton is at the end of the sixth β-strand, and the Glu that facilitates departure of the 3-OH group is in a loop at the end of the second β-strand.

(2) In the β-methylaspartate ammonia lyase (MAL) subgroup (~20 sequences), the Lys that initiates the reaction is at the end of the sixth β-strand but departure of 3-ammonium group does not need acid catalysis so no acidic group is located on the opposite face of the active site.

(3) In the muconate lactonizing enzyme (MLE) subgroup (~800 sequences), Lys acid/base catalysts are located on opposite faces of the active site at the ends of the second and sixth β-strands so that proton transfer from the substrate and to the enolate anion intermediate can occur from either face of the substrate/intermediate.

(4) In the mandelate racemase (MR) subgroup (~1100 sequences), a His-Asp dyad acid/base catalyst is at the ends of the seventh and sixth β-strands, respectively; both structures and sequence analyses reveal the presence of additional, variable acid/base catalysts at the ends of the second, third, or fifth β-strands.

(5) In the D-glucarate dehydratase (GlucD) subgroup (~130 sequences), a His-Asp dyad acid/base catalyst is also at the ends of the seventh and sixth β-strands, and a Lys acid/base catalyst is at the end of the second β-strand; however, an Asn ligand for the Mg^{2+} is at the end of the fifth β-strand.

To date, 15 reactions have been associated with members of the enolase superfamily (Figure 2): (1) enolase subgroup, the ubiquitous enolase reaction (*10*, *11*); (2) MAL subgroup, the MAL reaction (*12*, *13*); (3) MLE subgroup (framed in red), the MLE (*14*), *o*-succinylbenozate synthase (*15*), L-Ala-D/L-Glu epimerase (*16*, *17*), and N-succinylamino acid racemase reactions (*18*−*20*); and (4) MR subgroup (framed in blue), the MR (*21*), D-galactonate dehydratase (GalD) (*4*), D-gluconate dehydratase reactions (*2*, *3*); L-fuconate dehydratase (FucD) (*22*), D-tartrate dehydratase (TarD) (*23*), and L-talarate/galactarate dehydratase (TalrD/GalrD) (*24*); (5) GlucD subgroup, the D-glucarate/L-idarate dehydratase (GlucD) (*25*, *26*). However, on the basis of phylogenetic and network-based analysis, at least one-half of the sequences within the superfamily have unknown functions; these are found in each of the five subgroups. Therefore, we are continuing to devise approaches for discovering these functions (*27*).

Although only three members of the MR subgroup catalyze the MR reaction (~100 sequences), with the remaining functionally assigned members catalyzing the dehydration of acid sugars (~ 385 proteins), the reaction catalyzed by mandelate racemase (MR) has served as the paradigm for understanding the structural bases for the abstraction of the α-proton of the substrate and subsequent partitioning of the resulting enediolate intermediate to product. In the MR-catalyzed reaction (*28*), acid/base chemistry occurs on opposite faces of the active site so that the 1,1-proton-transfer reaction can be catalyzed (Figure 1). Lys 166 at the end of the second β-strand of the barrel domain is the (*S*)-specific acid/base catalyst, and His 297 at the end of the seventh β-strand, hydrogen-bonded to Asp 270 at the end of the sixth β-strand, is the (*R*)-specific base. One carboxylate oxygen of the substrate is hydrogen-bonded to Glu 317 at the end of the eighth β-strand; the other carboxylate oxygen is coordinated to the Mg^{2+} as well as hydrogen-bonded to Lys 164 at the end of the second β-strand (in a KxK motif). As the α-proton is abstracted and the enolate intermediate is formed, the strengths of these interactions increase, thereby allowing the intermediate to be transiently stabilized and kinetically competent.

The same active site motif found in the active site of MR (three carboxylate ligands for the Mg^{2+}, a His-Asp dyad, a Lys at the end of the second β-strand, and Glu/Lys electrophilic catalysts) is used in the anti-dehydration reactions catalyzed by L-fuconate dehydratase (FucD; (*22*)), D-tartrate dehydratase (TarD; (*23*)), and L-talarate/galactarate dehydratase (TalrD/GalrD; (*24*)). In these anti-dehydration reactions, a Lys at the end of the second β-strand abstracts the 2-proton, and the His at the end of the seventh β-strand (in the His-Asp dyad) facilitates departure of the 3-OH group.

[2] In our previous publications on the enolase superfamily, we discussed three major subgroups (enolase, MLE, and MR). However, in this article we explicitly include (1) the small MAL subgroup (~20 sequences[3]) that contains orthologous MALs as well as MAL-related proteins of unknown function, and (2) the GlucD subgroup (~130 sequences) that is similar to the MR subgroup (previously included in the MR subgroup) except that the ligand for the Mg^{2+} at the end of the fifth β-strand of the barrel domain is an Asn instead of a Glu and an Asp instead of a Glu is located at the end of the eighth β-strand of the barrel domain. These differences apparently are responsible for an altered coordination geometry of the substrate to the Mg^{2+} (in the GlucD subgroup the carboxylate group of the substrate is a bidentate ligand of the Mg^{2+}; in the MR subgroup, one carboxylate oxygen and the 2-OH group of the substrate are ligands of the Mg^{2+}, and the second carboxylate oxygen is hydrogen-bonded to the Glu at the end of the eighth β-strand).

[3] http://sfld.rbvi.ucsf.edu/ (*1*).

Table 1: Conserved Residues in the Subgroups of the Enolase Superfamily

strand[a] subgroup	β2	β3	β4	β5	β6	β7	β8
Enolase	E	D	E	D	K	H	K
MAL	H	D	E	D	K		
MLE	KxK	D	E	D	K	-	E/DXD/G
MR	Kx(K,R,H,D,Y,x)	D	E	E	D	H	E
GlucD	KxK	D	E	N	D	H	D
ManD	R	D	E	E	R	H	E

[a] Location in the $(\beta/\alpha)_7\beta$-domain. No conserved residue is located at the end of the first β-strand.

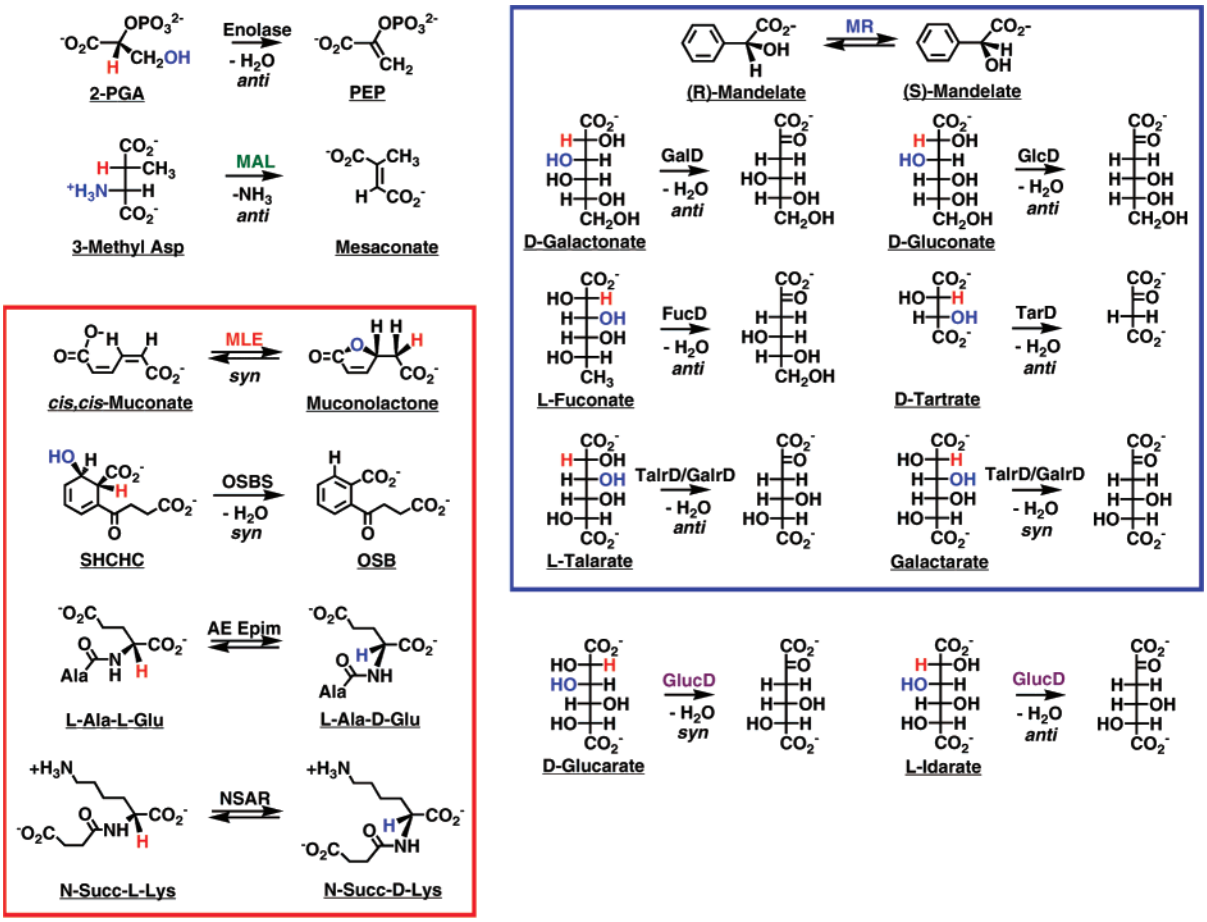
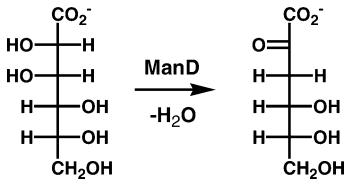


FIGURE 2: Reactions catalyzed by members of the enolase superfamily.

Although acid-catalyzed departure of the 3-OH group in the dehydration reactions catalyzed by FucD, TarD, and TalrD/GalrD and proton transfer to/from the α-carbon in the 1,1-proton-transfer reaction catalyzed by MR might be expected to have different geometric requirements, the Lys and His-Asp dyad acid/base catalysts in the active sites of MR, FucD, TarD, and TalrD/GalrD are remarkably superimposable. As might be expected from this conserved geometry, FucD and TalrD/GalrD also catalyze 1,1-proton transfer (epimerization) reactions in competition with their physiologically relevant dehydration reactions.

This manuscript focuses on the functional assignment and structural elucidation of members of a large, divergent orthologous group (family) of acid sugar dehydratases. We used a library of acid sugars to discover that the member encoded by the genome of *Novosphingobium aromaticivorans* catalyzes only the syn-dehydration of D-mannonate (the D-mannonate dehydratase or ManD reaction). Structures of both unliganded ManD as well as substrate/product

Scheme 1



liganded complexes reveal novel identities and locations for the acid/base catalysts: (1) the Tyr 159 base (in a hydrogen-bonded dyad with Arg 147 at the end of the second β-strand) that initiates the reaction is located in a loop of approximately 30 residues at the end of the second β-strand; (2) the acid that facilitates syn-elimination of the OH leaving group is either Tyr 159 or His 212 at the end of the third β-strand; (3) although a "spectator" His 312 is at the end of the seventh β-strand, no Asp hydrogen-bonding partner is at the end of the sixth β-strand. In addition, the carboxylate group of the substrate participates in electrophilic interactions with Arg 283 at the end of the sixth β-strand instead of a Lys at the

end of the second β-strand as in the MR subgroup (Table 1). Using sequence and network-based analysis (data not shown), all of the sequences that cluster with the ManD from *N. aromaticivorans* contain an inserted loop of ~30 residues at the end of the second β-strand that includes Tyr 159.[4] Thus, we conclude that ManD represents a sixth subgroup of the enolase superfamily, establishing the structural bases for additional functional diversity in catalyzing the dehydration of acid sugars.

MATERIALS AND METHODS

^{1}H NMR spectra were recorded on a Varian Unity 500NB MHz spectrometer. All compounds used were of the highest available commercial grade.

Cloning, Expression, and Purification of GI:146275900 from Novosphingobium aromaticivorans. The gene encoding GI:146275900 was PCR amplified from *N. aromaticivorans* genomic DNA (generous gift from Dr. David Balkwill, Florida State University) using platinum *Pfx* DNA polymerase (Invitrogen). The PCR reaction (100 μL) contained 1 ng of DNA, 10 μL of 10× *Pfx* amplification buffer, 1 mM $MgSO_4$, 0.4 mM dNTP, 40 pmol of each primer, and 1 unit of platinum *Pfx* polymerase. The gene was amplified on a PTC-200 gradient thermal cycler (MJ Research) with the following parameters: 94 °C for 2 min followed by 40 cycles of 94 °C for 1 min, a gradient of 45 °C to 65 °C for 1 min and 15 s, 68 °C for 2 min followed by a final 10 min extension at 68 °C. The amplified gene was cloned into the pET 15b vector (Novagen).

All proteins were expressed in *Escherichia coli* strain BL21 (DE3). Transformed cells were grown at 37 °C for 48 h in Luria−Bertani broth (LB) supplemented with 100 μg/mL ampicillin and then harvested by centrifugation. No IPTG induction was used. The cells were resuspended in binding buffer (20 mM Tris-HCl, pH 7.9, 5 mM $MgCl_2$, 500 mM NaCl, 5 mM imidazole) and lysed by sonication. The lysate was clarified by centrifugation (10 min at 4500*g*), and the supernatant was fractionated on a chelating Sepharose Fast Flow (American Biosystems) column charged with Ni^{2+}. The cell lysate was applied, and the His-tagged protein was eluted with a gradient of imidazole (0 to 0.5 M). The N-terminal His-tag could not be cleaved by thrombin.

Site-Directed Mutagenesis. All mutants were constructed using overlap extension mutagenesis. The PCR reactions (50 μL) to construct megaprimers contained 1 ng of the plasmid encoding wild type ManD in the pET17b vector, 5 μL of 10× PCR buffer, 4 mM $MgCl_2$, 2 mM dNTPs, 40 pmol of each primer, 1 unit of *Taq* DNA polymerase (Invitrogen), and 0.5 units of cloned *Pfu* polymerase (Stratagene). The 5′-megaprimer was constructed with the T7pro primer paired with a 3′-mutant primer; the 3′-megaprimer was constructed with the T7term primer and a 5′-mutant primer. The PCR parameters: 95 °C for 4 min followed by 26 cycles of 95 °C for 45 s, 55 °C for 45 s, and 72 °C for 2 min and 15 s followed by 7 min at 72 °C. The megaprimers were purified by 1% agarose gel electrophoresis. The second PCR reaction (50 μL) contained 5 μL 10× PCR buffer, 4 mM $MgCl_2$, 2 mM dNTPs, 40 pmol of T7pro and T7term primers, 200 pmol of each of the megaprimers, 1 unit of *Taq* DNA polymerase (Invitrogen), and 0.5 units of cloned *Pfu* polymerase (Stratagene). The same PCR program was used.

Cloning, Expression, and Purification of 2-Keto-3-deoxygluconate Kinase (KdgK). The gene encoding 2-keto-3-deoxy-D-gluconate kinase (KdgK) was amplified from *E. coli* K12 genomic DNA (ATCC 700928D) using the PCR procedures described above and ligated into pET 17b. The cell lysate was purified by chromatography on a Dowex DEAE column (equilibrated with 10 mM Tris-HCl, pH 7.9, and 5 mM $MgCl_2$, using a gradient from 0 to 1 M NaCl. Following analysis by SDS-PAGE, fractions were combined, concentrated with an Amicon 200 mL concentrator, dialyzed into 10 mM, Tris-HCl pH 7.9, and 5 mM $MgCl_2$ and chromatographed on a Hi-Performance Q-Sepharose column using a gradient of 0 to 1 M NaCl. Fractions containing protein were analyzed by SDS-PAGE, concentrated, and dialyzed into 10 mM Tris-HCl, pH 7.9, containing 5 mM $MgCl_2$, and 100 mM NaCl.

Screen for Dehydration Activity. GI:146275900 was screened for dehydration activity using a library of mono- and diacid sugars as described previously (*22*−*24*). Briefly, the enzyme (1 μM) was incubated individually with the members of a library of acid sugars overnight at 30 °C in 50 mM Tris, pH 7.9, containing 5 mM $MgCl_2$ and 10 mM acid sugar. The reactions (200 μL) were quenched with a 5-fold volume excess of 1% sodium acetate/1% semicarbazide-HCl and incubated for 1 h at room temperature before the absorbance at 250 nm (ϵ = 10 200 M^{-1} cm^{-1}) was recorded.

Stereochemical Assignment of Resonances Associated with the C3 Protons of the Dehydration Product. D-Mannonate was dehydrated in a reaction (800 μL, 25 °C) containing 10 mM D-mannonate, 50 mM potassium phosphate, pH 7.9, 5 mM $MgCl_2$, and 1 μM enzyme. The reaction was incubated at 30 °C for 16 h before lyophilization and resuspension in 800 μL of D_2O.

The product was analyzed by COSY to determine which sets of 3-pro*R* and 3-pro*S* resonances were coupled (a mixture of anomeric β-pyranose and β-furanose hemiacetals). The geminal and vicinal ^{1}H−^{1}H coupling constants [(*29*); L. Williams and F. M. Raushel, personal communication] were confirmed by 1D NOE experiments in which the resonances associated with H4 of the β-pyranose and β-furanose forms were irradiated and magnetization transfer to H3 was quantitated.

Stereochemical Assignment of Proton Incorporation. ManD was exchanged into D_2O buffer by dilution of 1 mL of 100 μM ManD with 10 mL of D_2O and concentration to 1 mL in a 10 mL Amicon filter with a Millipore 10 000 MW polyethersulfone ultrafiltration membrane; this procedure was repeated twice. A reaction (800 μL, 30 °C) was then performed with 10 mM D-mannonate in 50 mM Tris-d_{11}-DCl, pD 7.9, containing 5 mM $MgCl_2$ and 1 μM ManD.

Kinetic Assay of D-Mannonate Dehydratase (ManD). ManD activity was quantitated using either a discontinuous assay utilizing semicarbazide (*30*) or a continuous, coupled-enzyme spectrophotometric assay. In the latter assay, the product was 6-phosphorylated using KdgK and ATP, and formation of ADP was measured using pyruvate kinase (PK) and L-lactate dehydrogenase (LDH). The assay (1 mL, 25

[4] A few other sequences in the MR subgroup, including some tentatively identified as D-gluconate dehydratases (*2*, *3*), also have insertions in this approximate position, but these are shorter than those found in the ManD family.

Table 2: Data Collection and Refinement Statistics

	SeMet-ManD·Mg	K271E ManD·Mg·DMN[a]	ManD·Mg·DDG[b]
	Data Collection		
beamline	NSLS X4A	NSLS X4A	NSLS X4A
wavelength (Å)	0.97915	0.979	0.979
space group	$C222_1$	$C222_1$	$C222_1$
no. of molecules in a.u.	4	4	4
	Unit Cell Parameters		
a (Å)	117.465	116.254	116.269
b (Å)	167.711	165.312	165.241
c (Å)	166.441	166.853	167.039
resolution (Å)	1.8	2.2	2.0
no. of unique reflections	146496	78216	104980
completeness (%)	96.8	96.1	96.9
R_{merge}	0.086	0.061	0.073
average *I*/*s*	30.4	35.9	34.5
	Refinement		
resolution	25.0−1.8	25.0−2.2	25.0−2.0
R_{cryst}	0.178	0.168	0.162
R_{free}	0.192	0.206	0.192
rmsd for bonds (Å)	0.006	0.005	0.005
rmsd for angles (deg)	1.4	1.3	1.3
no. of protein atoms	12512	12222	12227
no. of waters	858	647	850
no of Mg^{2+} ions	4	4	4
bound ligand	-	MNN	DDG
ligand atoms	-	56	52
PDB entry	2QJJ	2QJM	2QJN

[a] D-Mannonate. [b] 2-Keto-3-deoxy-D-gluconate.

°C) contained 50 mM potassium HEPES, pH 7.5, 5 mM $MgCl_2$, 1.5 mM ATP, 1.5 mM PEP, 0.16 mM NADH, 9 units of PK, 9 units of LDH, 18 units of KdgK, and ManD (wild type or mutant). Dehydration was quantitated by measuring the decrease in absorbance at 340 nm ($\epsilon = 6220$ M^{-1} cm^{-1}).

Crystallization and Refinement. Three different crystal forms (Table 2) were grown by the hanging drop method at room temperature: (1) selenomethionine (SeMet)-substituted ManD liganded with Mg^{2+}, (2) the K271E mutant of ManD liganded with Mg^{2+} and D-mannonate, (3) wild type ManD liganded with Mg^{2+} and 2-keto-3-deoxy-D-gluconate. The crystallization conditions utilized the following conditions:

(1) For SeMet-substituted wild type ManD liganded with Mg^{2+}, the protein solution contained SeMet-substituted wild type ManD from *N. aromaticivorans* (25 mg/mL) in 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM β-mercaptoethanol, and 5 mM $MgCl_2$; the precipitant contained 12% PEG-3350, 0.2 M ammonium sulfate, and 0.1 M Tris-HCl (pH 8.5). Crystals appeared in 3 days and exhibited diffraction consistent with space group $C222_1$, with four molecules of ManD per asymmetric unit.

(2) For the K271E mutant of ManD liganded with Mg^{2+} and D-mannonate, the protein solution contained the K271E mutant of ManD (9.4 mg/mL) in 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, and 40 mM D-mannonate; the precipitant contained 2.1 M NaCl, and 0.1 M Bis Tris, pH 5.5. Crystals appeared in 5 days and exhibited a diffraction pattern consistent with space group $C222_1$ with four molecules of the mutant ManD per asymmetric unit.

(3) For wild type ManD liganded with Mg^{2+} and 2-keto-3-deoxy-D-gluconate, the protein solution contained wild type ManD (37.7 mg/mL) in 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, and 40 mM D-mannonate. The precipitant contained 8% PEG 3350, 0.1 M Tris, pH 8.5, and 0.2 M ammonium sulfate. Crystals appeared in 6 days and exhibited diffraction consistent with space group $C222_1$ with four molecules of ManD per asymmetric unit.

Prior to data collection, the crystals were transferred to cryoprotectant solutions composed of their mother liquids and 20% glycerol. After incubation for ~15 s, the crystals were flash-cooled in a nitrogen stream. A one-wavelength single anomalous dispersion (SAD) data set for a crystal of SeMet-substituted ManD (Table 2, column 1) was collected to 1.8 Å resolution at the NSLS X4A beamline (Brookhaven National Laboratory) on an ADSC CCD detector. Data sets for complexes of the K271E mutant with Mg^{2+} and D-mannonate (column 2), and wild type with Mg^{2+} and 2-keto-3-deoxy-D-gluconate (column 3), were collected at the same beamline to the resolution 2.2 Å and 2.0 Å, respectively. Diffraction intensities were integrated and scaled with DENZO and SCALEPACK (*31*). The data collection statistics are given in Table 2.

Structure Determination, and Refinement. The structure of the SeMet-substituted ManD was determined by SAD with SOLVE (*32*); 45 of the 48 selenium sites were identified. These heavy atom sites were used to calculate initial phases which were improved by solvent flattening and NCS-averaging with RESOLVE (*33*), yielding an interpretable map for four monomers in the asymmetric unit for space group $C222_1$. Iterative cycles of automatic rebuilding with ARP (*34*), manual rebuilding with TOM (*35*), and refinement with CNS (*36*) resulted in a model at 1.8 Å with R_{cryst} of 0.178 and an R_{free} of 0.192. The residues 156−172 in monomers C and D are disordered and are not included in the final model. The Mg^{2+} ions are clearly visible in the active sites of all four protein monomers.

The structure of the K271E mutant of ManD crystallized with Mg^{2+} and D-mannonate was determined by molecular replacement with PHASER (*37*), using the SeMet-substituted

ManD structure as the search model. Iterative cycles of automatic rebuilding with ARP, manual rebuilding with TOM, and refinement with CNS were performed. The model was refined at 2.2 Å with an R_{cryst} of 0.168 and an R_{free} of 0.206. The Mg^{2+} ions were clearly visible in the electron density maps for all four protein molecules in the asymmetric unit. The molecules of D-mannonate were well-defined in each polypeptide.

The structure of the wild type ManD cocrystallized with Mg^{2+} and D-mannonate was also determined by molecular replacement using the SeMet-substituted ManD structure as a search model. The model was refined at 2.0 Å with an R_{cryst} of 0.162 and an R_{free} of 0.192. The Mg^{2+} ions were well defined in all four polypeptides in the asymmetric unit. The electron density for the ligand bound in the active site unambiguously corresponded to the product 2-keto-3-deoxy-D-gluconate despite the fact that the initial protein solutions contained the D-mannonate substrate. This structure and the structure of the K271E mutant of ManD liganded with Mg^{2+} and D-mannonate have chain segments 156−172 disordered in all four monomers in both structures. These chain segments belong to the loop located between second β-strand and second α-helix of the barrel. The disordered segments were not included in the final models for both complexes. Final refinement statistics for all three structures are given in Table 2.

RESULTS AND DISCUSSION

The genome of *E. coli* K-12 encodes eight members of the enolase superfamily, five of known function [enolase (GI:16130686), *o*-succinylbenozate synthase (GI:16130196), L-Ala-D/L-Glu epimerase (GI:90111249), D-glucarate/L-idarate dehydratase (GlucD; GI:16130695), D-galactonate dehydratase (GalD; GI:49176390)] and three of unknown function [RspA (GI:16129539), YfaW (GI:16130182), and a GlucD-related protein (GlucDRP; GI:16130694)].[3] We would like to complete the functional assignments of the *E. coli* K-12-encoded members of the enolase superfamily; this manuscript focuses on the functional assignment of RspA. Therefore, our preliminary investigations focused on the RspA from *E. coli* K-12, but this protein could not be purified as a result of unavoidable proteolytic digestion during purification. As a result, we expressed and purified several orthologues, with that from *N. aromaticivorans* DSM 12444 (GI:146275900) as the focus of this manuscript.

As of July 20, 2007, the sequence databases contained 71 homologues of the RspAs from *E. coli* and *N. aromaticivorans* that share the active site functional groups identified and structurally characterized in the present study;[3] 60 are expected to catalyze the same reaction using the same substrate as RspA (orthologues) based on conservation of the acid/base and specificity-determining residues identified in the present study.

A structure-informed sequence alignment of the RspAs with other members of the enolase superfamily (Figure 3) discloses the identities of the ligands for the essential Mg^{2+}, Asp 210, Glu 236, and Glu 262, at the ends of the third, fourth, and fifth β-strands of the barrel domain, respectively (*N. aromaticivorans* RspA numbering), as well as the electrophilic Glu 339 at the end of the eighth β-strand. Although a conserved His 312 aligns with the His of the conserved His-Asp dyad that is characteristic of the MR subgroup, neither the sequence alignment nor a structural alignment with the structurally characterized members of the MR subgroup reveals an Asp that aligns with the Asp of the conserved His-Asp dyad found in the MR subgroup. In addition, the alignment does not reveal a candidate for either an acid/base catalyst or an electrophilic Lys at the end of the second β-strand; instead, the alignment reveals a conserved Arg 147-x-Gln 149 motif at a position consistent with a location at the end of the second β-strand.

Identification of D-*Mannonate as the Substrate for RspA from N. aromatoicivorans (GI:146275900*). We previously described the use of a library of mono- and diacid sugars to discover the TarD (*23*) and TalrD/GalrD (*24*) functions in the MR subgroup. With that library, we discovered that only D-mannonate is a substrate for the homologue of RspA encoded by the genome of *N. aromaticivorans*. The dehydration was confirmed by ^1H NMR spectroscopy (Figure 4, panel B); the product (2-keto-3-deoxy-D-gluconate) is a mixture of four species, anomeric pairs of both furanose and pyranose hemiacetals (Figure 4, panel A).

The kinetic constants were measured using a coupled-enzyme, spectrophotometric assay. The values of k_{cat} and k_{cat}/K_m, 1.3 ± 0.1 s^{-1} and 6.9×10^3 M^{-1} s^{-1}, respectively, are comparable to those measured for other reactions catalyzed by acid sugar dehydratases in the enolase superfamily, supporting the conclusion that the physiological function of the RspA from *N. aromaticivorans* is the D-mannonate dehydratase (ManD) reaction.

Further support for this conclusion is obtained from genome context in the *N. aromaticivorans* genome:[5] in one direction, a gene encoding KdgK is one removed from the gene encoding ManD, and, in the other direction, genes encode orthologues of mannitol dehydrogenase and glucuronate isomerase. Like analogous ManDs that are members of a distinct Fe^{2+}-dependent superfamily of acid sugar dehydratases, KdgK and uronate isomerase are in the glucuronate catabolic pathway, so the genome context provides additional evidence that the ManD function we discovered by library screening is physiologically relevant. Given the levels of sequence identity (>50%) shared with homologues that have the same acid/base catalysts and specificity-determining residues, we assign the ManD function to these proteins, including the RspA protein encoded by the *E. coli* K-12 genome.[6]

Structure of ManD. Three structures were solved for the ManD from *N. aromaticivorans*: (1) wild type protein at pH 8.5 in the presence of Mg^{2+}; (2) an accidental K271E mutant at pH 5.5 in the presence of Mg^{2+} and the D-mannonate substrate; (3) wild type protein at pH 8.5 in the presence of Mg^{2+} and the 2-keto-3-deoxy-D-gluconate dehydration product. The K271E mutant is active at neutral pH,[7] so the presence of the substrate in its active site likely is a consequence of the acidic conditions adventitiously required for its crystallization. Representative electron

[5] http://img.jgi.doe.gov/cgi-bin/pub/main.cgi

[6] The genome of *E. coli* K-12 also encodes the analogous Fe^{2+}-dependent D-mannonate dehydratase (UxuA) in the D-glucuronate catabolic pathway. The physiological function of the ManD discovered in this study is uncertain.

[7] $k_{cat} = 2.7$ s^{-1}, $k_{cat}/K_m = 4 \times 10^3$ M^{-1} s^{-1}.

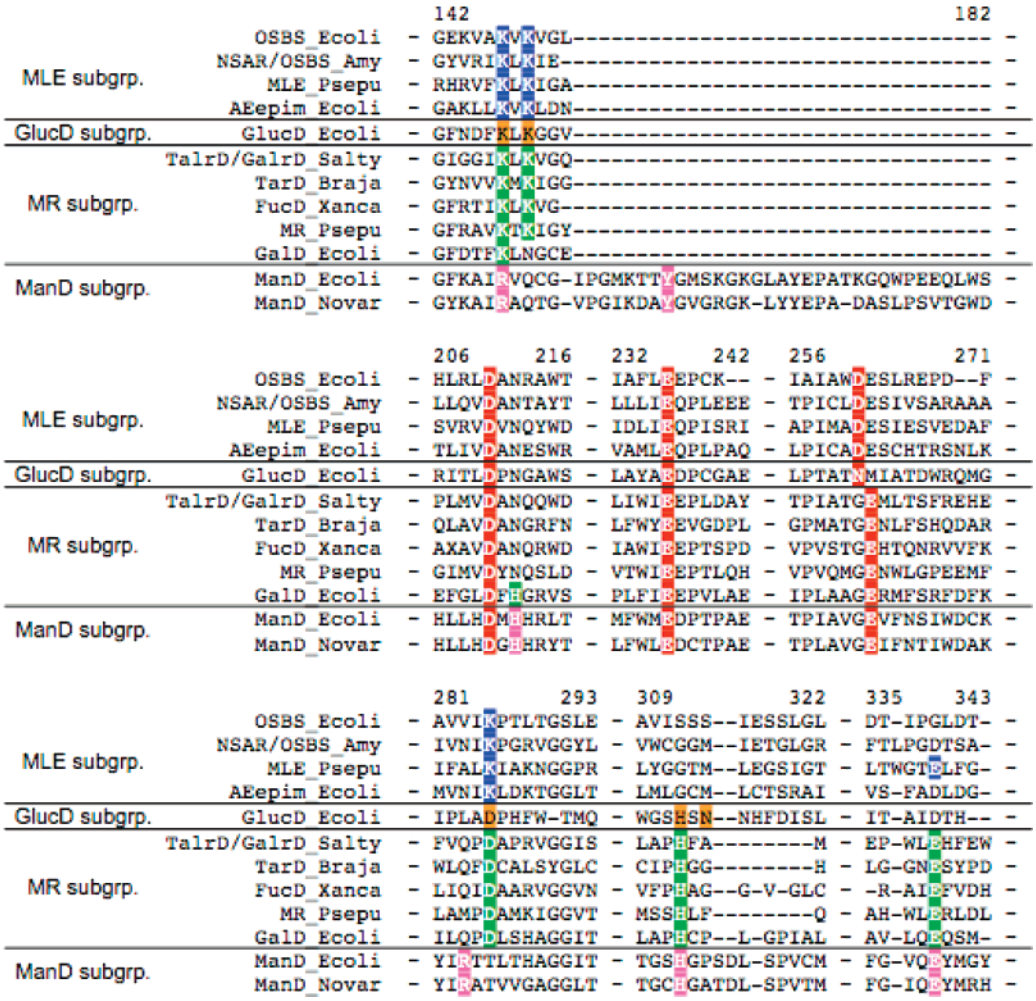


FIGURE 3: Sequence alignment of selected members of the MLE, GlucD, MR, and ManD subgroups of the enolase superfamily. The conserved metal ligands common to all members of the enolase superfamily are highlighted in red. Residues required for catalysis in the MLE, GlucD, and MR subgroups are highlighted in blue, orange, and green, respectively. Residues identified in this study as important for catalysis by ManD are highlighted in magenta. Residue numbering is relative to *N. aromaticivorans* ManD. Tildes (∼) indicate regions in the alignment that have been excluded for brevity. Sequences are *o*-succinylbenzoate synthase from *E. coli* (OSBS_Ecoli), promiscuous *N*-succinylamino acid racemase/*o*-succinylbenzoate synthase from *Amycolatopsis* sp. T-1−60 (NSAR/OSBS_Amy), MLE from *Pseudomonas putida* (MLE_Psepu), L-Ala-D/L-Glu epimerase from *E. coli* (AEepim_Ecoli), GlucD from *E. coli* (GlucD_Ecoli), TalrD/GalrD from *Salmonella typhimurium* LT2 (TalrD/GalrD_Salty), TarD from *Bradyrhizobium japonicum* (TarD_Braja), FucD from *Xanthomonas campestris* (FucD_Xanca), MR from *P. putida* (MR_Psepu), GalD from *E. coli* (GalD_Ecoli), ManD from *E. coli* (ManD_Ecoli), and ManD from *N. aromaticivorans* (ManD_Novar).

density for the active site region in structure 2 is shown in Figure 5.

The protein (402 residues) is organized in two domains, as previously observed for all other structurally characterized members of the enolase superfamily (Figure 6, left structure): an N-terminal $\alpha+\beta$ domain that extends from residue 1−118, and a $(\beta/\alpha)_7\beta$-barrel domain that extends from residue 119−375. Residues 376−402 are at the interface between the domains, with the carboxylate group of the C-terminal Trp 402 participating in a hydrogen-bond interaction with the active site His 212 at the end of the third β-strand.

The minimum biological unit is a dimer because the active site is at the interface of two polypeptides and includes residues from both polypeptides (*vide infra*); in solution, ManD is a tetramer as judged by gel filtration (data not shown). In MR, FucD, and TarD, the second polypeptide in a dimer provides a single side chain that participates in the binding pocket for the distal portion of the substrate; in TalrD/GalrD the active site is formed from a single polypeptide. In ManD, the OH group of Tyr 75 in the adjacent polypeptide is hydrogen-bonded to the carboxylate oxygen of Glu 262, the ligand for the Mg^{2+} at the end of the fifth β-strand; this OH group is also proximal to the imidazole group of His 213 that follows Asp 210, the ligand for the Mg^{2+} at the end of the third β-strand (Figure 7, panel B). Also, Trp 76 from the adjacent polypeptide participates in the binding pocket for C6 of the substrate/product (Figure 7, panel C).

The structures of MR, a representative structurally characterized member of the MR subgroup, and ManD are compared in Figure 6. Although these can be well-superimposed (rmsd is 1.24 Å for 192 pairs of α-carbons) and the structures of the N-terminal $\alpha+\beta$ capping and the $(\beta/\alpha)_7\beta$-barrel domains are homologous, the two structures are readily distinguished by the identities of the secondary structural elements that sequester the substrate/product from the solvent: in ManD, these are the Val 152 − Thr 183 loop at the end of the second β-strand of the barrel domain and the C-terminus (Pro 383 − Trp 402); in MR, Pro 17−Gly

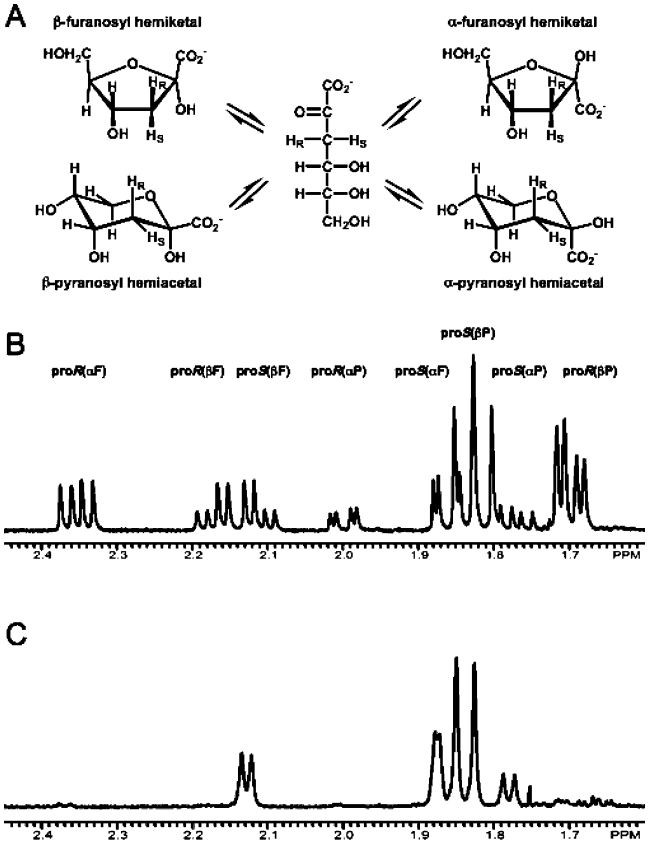


FIGURE 4: Panel A, the structures of the acyclic keto and cyclic furanose and pyranose forms of the 2-keto-3-deoxy-D-gluconate product. Panel B, partial ^{1}H NMR spectrum of the product obtained by a reaction in H_2O showing the resonances associated with the 3-pro*R* and 3-pro*S* protons. Panel C, partial ^{1}H NMR spectrum of the product obtained by a reaction in D_2O.

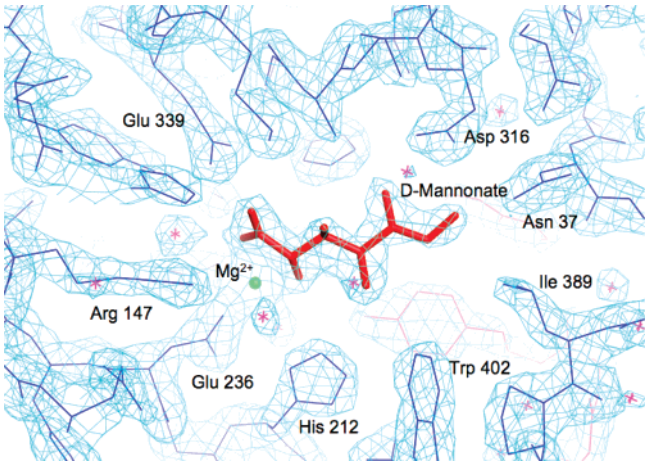


FIGURE 5: Representative electron density for the active site of the K271E mutant of ManD liganded with D-mannonate and Mg^{2+} contoured at 1.5 σ. The details of the interactions between D-mannonate and the active site are described in the text.

30 in the 20s loop of the $\alpha+\beta$ capping domain serve this function. In ManD the 20s loop is reduced to a turn. The Val 152 − Thr 183 loop contains Tyr 159, the likely base that abstracts the 2-proton to initiate the reaction (*vide infra*).

As expected from the sequence alignment, Asp 210, Glu 236, and Glu 262, at the ends of the third, fourth, and fifth β-strands of the barrel domain, respectively, are the ligands for the Mg^{2+} (Figure 7, panel B). Also, while His 312 is at the end of the seventh β-strand, it is not located in a His-Asp dyad as found in members of the MR subgroup; instead, the imidazole nitrogens of His 312 are hydrogen-bonded to the electrophilic Glu 339 at the end of the eighth β-strand and O5 of the bound substrate/product.

The absence of the His-Asp dyad at the ends of the seventh and sixth β-strands (i.e., absence of a homologue of the Asp) as well as the identification and location of Tyr 159 as the base that initiates the reaction distinguish ManD from the members of the MR subgroup and establish the identity of the ManD subgroup, the sixth in the enolase superfamily (Table 1).

The Active Site of ManD. Because the Val 152−Thr 183 loop that sequesters the substrate/product from solvent is ordered in the absence of ligands but disordered in both the substrate- and product-liganded structures, the expected structure of a substrate-liganded complex was constructed by superimposing the structures of the unliganded and substrate-liganded proteins; the α-carbons superimpose with an rmsd of 0.2 Å for 384 pairs of α-carbons (Figure 7, panel A). The composite active site is shown in Figure 7, panels B and C (unliganded in dark gray; substrate liganded in dark cyan, with D-mannonate in gold). The superpositioning of of these residues illustrates the minimal effect of ligand binding on the structure of the active site (with the exception of the Val 152−Thr 183 loop).

The candidates for the residues that are directly involved in catalysis (ligands for the Mg^{2+} and both electrophilic and acid/base catalysts) are shown in Figure 7, panel B; these are conserved in all of the putative orthologues in the sequence databases. As observed in the active sites of MR, FucD, TarD, and TalrD/GalrD, one carboxylate oxygen of the D-mannonate substrate and its 2-OH group are bidentate ligands of the Mg^{2+}; the same carboxylate oxygen is hydrogen bonded to Arg 283 at the end of the sixth β-strand. This contrasts with the members of MLE, MR, and GlucD subgroups in which a Lys residue at the end of the second β-strand (the first in a conserved Lys-X-Lys motif) is the electrophilic residue that interacts with this carboxylate oxygen. The other carboxylate oxygen is hydrogen bonded to Glu 339, at the end of the eighth β-strand; a similar interaction is observed in the active sites of MR, FucD, TarD, and TalrD/GalrD. The difference in the positions and identities of the cationic electrophilic residue further distinguishes the ManD subgroup (Table 1).

The phenolic oxygen of Tyr 159 in the Val 152−Thr 183 loop is hydrogen-bonded to the guanidinium group of Arg 147, providing evidence that the oxygen is deprotonated. The oxygen is also positioned 2.4 Å from C2 of the D-mannonate substrate, providing structural evidence that it is the basic residue that initiates the reaction. The position and identity of this catalytic Tyr-Arg dyad contrasts with the Lys acid/base catalyst at the end of the second β-strand (the second in a conserved Lys-X-Lys motif) in the active sites of members of the MR, MLE, and GlucD subgroups of the superfamily.

These critical differences in active site structure are illustrated by the superposition of the catalytic groups in the active sites of MR (green) and ManD (gray) shown in Figure 8. The ligands for the essential Mg^{2+} are positionally conserved at the ends of the third, fourth, and fifth β-strands of the barrel domain (not shown in the figure); in both active sites one carboxylate oxygen and the α-OH group are

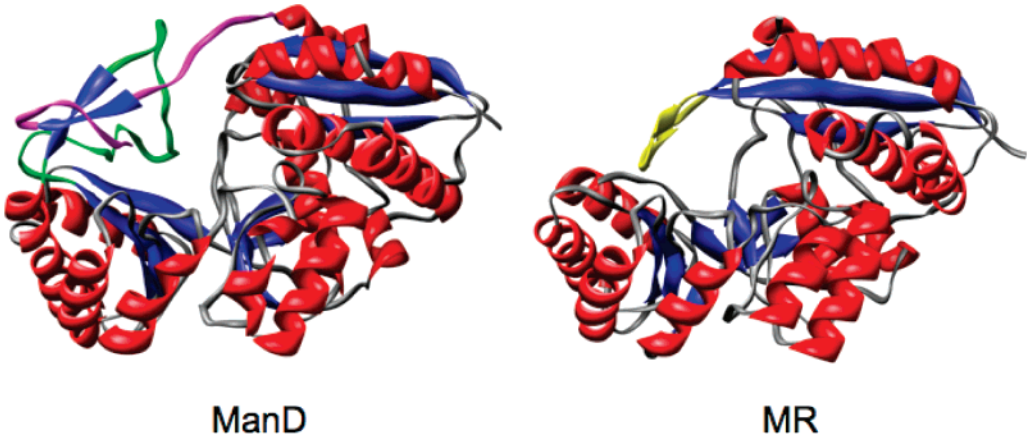


FIGURE 6: Structural comparison of ManD and MR. In ManD, the Val 152 − Thr 183 loop at the end of the second β-strand is highlighted in green and the Pro 383 − Trp 402 C-terminus is highlighted in magenta. In MR, the 20s loop in the C-terminal $\alpha+\beta$ capping domain is highlighted in gold.

bidentate ligands of the Mg^{2+}. In addition, in both active sites the second carboxylate oxygen is hydrogen bonded to the electrophlic Glu at the end of the eighth β-strand that is is also positionally conserved (not shown). However, in MR Lys 164 at the end of the second β-strand is coordinated to the substrate carboxylate oxygen that is coordinated to the Mg^{2+}; in ManD, the analogous electrophilic residue is Arg 283 located at the end of the sixth β-strand. In MR, Lys 166, also at the end of the second β-strand, is the S-specific base; in ManD, the Arg 147-Tyr 159 hydrogen-bonded dyad is the base. Note that the α-carbons of Lys 164 in MR and Arg 147 in ManD are structurally equivalent, but the cationic functional groups are located in different positions relative to the backbone and substrate so that these can serve different catalytic functions.

Abstraction of the 2-proton from the substrate will be accompanied by a change in hybridization of C2 from sp^3 to sp^2 as the enolate anion is formed. As a result of this change in geometry, the 3-OH group will move toward both Tyr 159 and His 212 at the end of the third β-strand; the protonated form of the latter residue is stabilized by a hydrogen bond with the carboxylate group of the C-terminal Trp 402. Either of these residues is positioned to function as the acid catalyst that facilitates vinylogous elimination of the 3-OH group in the syn-dehydration.

His 312, predicted and structurally confirmed to be at the end of the seventh β-strand, is positioned on the opposite face of the active site from the catalytic Tyr-Arg dyad, although it does not participate in a His-Asp dyad that is characteristic of members of the MR subgroup (Figure 8). Instead, it forms a hydrogen bond to O5 of the substrate, thereby assisting in determining substrate specificity.

The residues that form the binding pocket for the distal portion of the D-mannonate substrate are shown in Figure 7, panel C. Hydrogen bonding interactions involve O4 (His 212), O5 (His 312 and Asp 316), and O6 (Asn 37 and Asp 316); a hydrophobic surface of the binding pocket is provided by Leu 389 and Trp 402 in addition to Tyr 75 and Trp 76 from the adjacent polypeptide. These residues are conserved in at least 60 of the 71 homologous sequences identified as members of the ManD subgroup using the criteria described in this section; we assume that the sequences that share these specificity-determining residues are orthologues.

Mechanism of the ManD-Catalyzed Reaction: Site-Directed Mutagenesis of Active Site Residues. On the basis of the proximity and conservation of residues that are positioned to be candidates for direct involvement in catalysis, we constructed and characterized several mutants: Y159F (the putative acid/base catalyst), R147A and R147K (modulation of the pK_a of the phenolic OH group of Tyr 159), H212N (a possible acid catalyst), R283A (an electrophilic catalyst), H312N (the homologue of the His in the His-Asp dyad in the MR subgroup), and E339Q (an electrophilic catalyst). As judged by both the coupled-enzyme assay using KdgK and ^{1}H NMR analysis of reaction mixtures, only the R147K mutant had detectable catalytic activity ($k_{cat} = 0.28\ s^{-1}$ and $k_{cat}/K_m = 3.6 \times 10^1\ M^{-1}\ s^{-1}$).

The lack of activity associated with the Y159F substitution is consistent with its role as the base that initiates the reaction; the lack of activity associated with the R283A and E339Q substitutions is consistent with their roles as electrophilic catalysts that stabilize the enolate anion intermediate, based on the putative structure of the D-mannonate complex. The lack of activity associated with the H212N and H312N substitutions is consistent with their strict conservation in orthologous ManDs. Perhaps the phenotype associated with the H212N substitution is caused by its involvement in facilitating the departure of the 3-OH group, but a precise description of their mechanistic involvement is uncertain.[8] That the R147K substitution but not the R147A substitution is active is consistent with the hypothesis that Arg 147 modulates the pK_a of the phenolic OH group of Tyr 159, thereby enabling it to abstract the 2-proton.

Mechanism of the ManD-Catalyzed Reaction: Stereochemical Course of the Dehydration Reaction. The structure of the substrate-liganded complex establishes that ManD catalyzes a syn-dehydration. We determined the stereochemical course of replacement of the 3-OH group by solvent hydrogen by performing a reaction in a D_2O-containing solvent in order to specify the location and, therefore, identity of the acid catalyst that delivers a proton to the enolate intermediate derived from dehydration.

[8] In principle, the role of His 212 could be probed using the H212N mutant and 3-deoxy-3-fluoro-D-mannonate as an alternate substrate (*4*). However, a synthesis of this compound has not been reported.

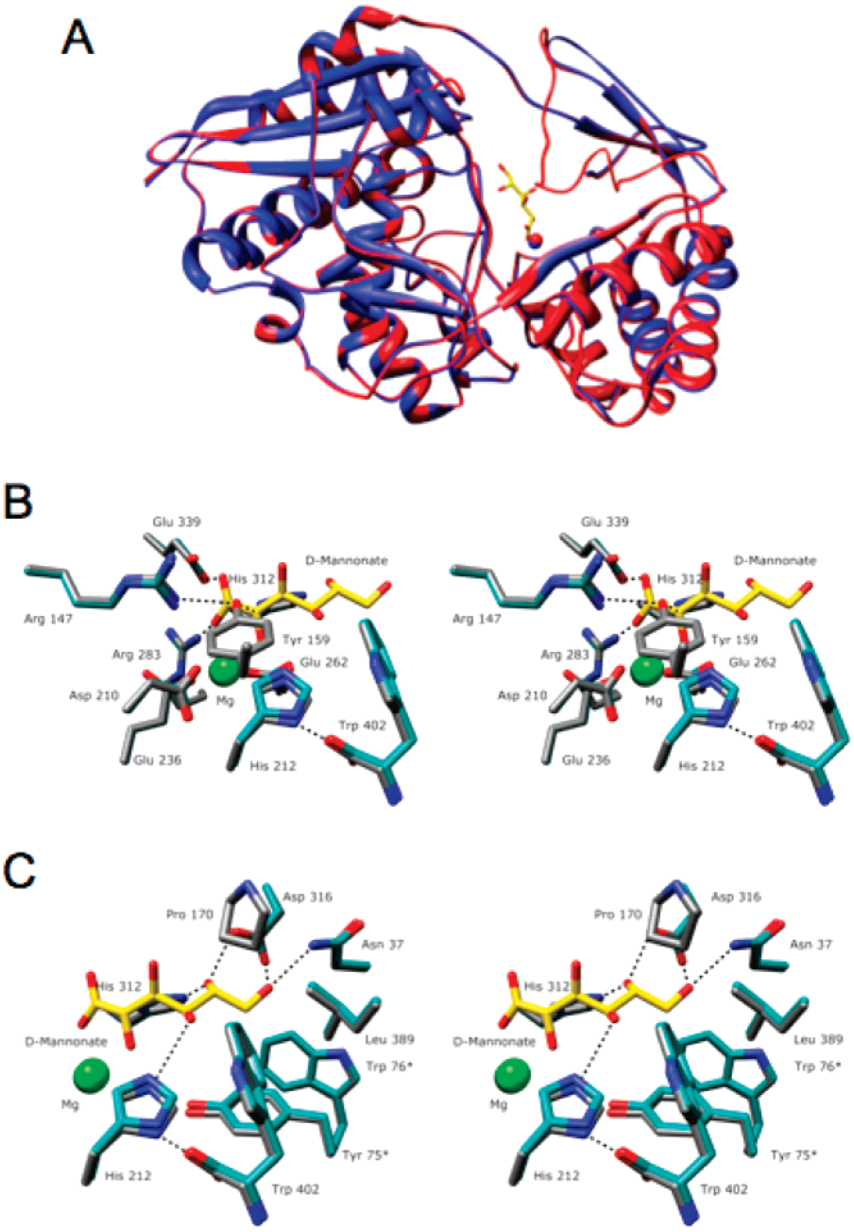


FIGURE 7: Panel A, superposition of the structures of the unliganded wild type protein (blue) and the K271E liganded with D-mannonate (red; D-mannonate is highlighted in yellow). Panel B, stereoview of the predicted active site of ManD, showing the residues likely involved in catalysis. Panel C, stereoview of the predicted active site of ManD, showing the residues involved in substrate specificity. In panels B and C, the wild type, unliganded structure is gray, and the K271E liganded structure is cyan.

As illustrated in Figure 4, panel A, the 2-keto-3-deoxy-D-gluconate product exists as a mixture of four species, pairs of anomers of both furanose and pyranose hemiacetals. The ^{1}H NMR resonances of the geminal 3-hydrogens of each hemiacetal are resolved and can be associated with one another by COSY spectroscopy (Figure 4, panel B). In studies of the stereochemical course of the reaction catalyzed by uronate isomerase, Raushel and co-workers assigned the chemical shifts of the resonances associated with each of the four species (*29*).

We used that information as well as the magnitudes of the vicinal C3−C4 ^{1}H−^{1}H coupling constants to assign the resonances of the 3-pro*S* and 3-pro*R* hydrogens. As shown in Figure 4, panel C, when the ManD-catalyzed reaction was performed in D_2O, the resonances associated with each of the 3-pro*R* protons were absent, and the resonances associated with the 3-pro*S* protons were simplified by loss of the geminal ^{1}H−^{1}H coupling. The loss of the 3-pro*R* resonances establishes that the replacement of the 3-OH group of the D-mannonate substrate with solvent hydrogen occurs with *retention* of configuration and, therefore, that *all* of the proton-transfer reactions involved in the syn-dehydration are enzyme-catalyzed and occur from the same face of the active site. As noted previously, Tyr 159 is positioned to initiate the reaction by abstraction of the 2-proton; both Tyr 159 and His 212 are positioned appropriately to catalyze both acid-catalyzed departure of the 3-OH group to generate an enol intermediate and its subsequent ketonization by protonation of carbon-3.

On the basis of the proposed structure of the D-mannonate-liganded complex and the stereochemical course of the replacement of the 3-OH group with solvent deuterium, we conclude that the mechanism of the ManD-catalyzed reaction is that shown in Figure 9.

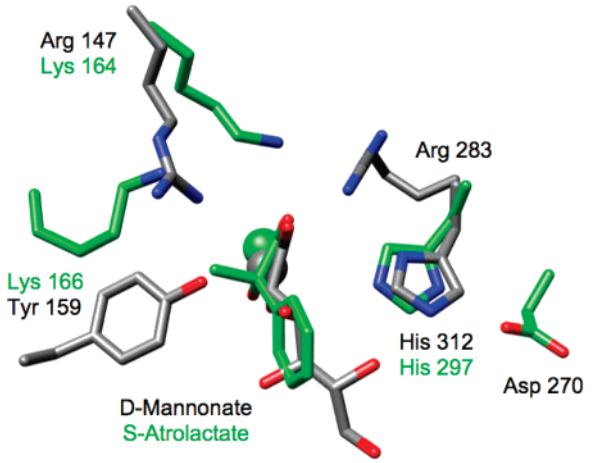


FIGURE 8: Comparison of the active sites of ManD (gray) and MR (green) illustrating differences in the catalytic residues. The essential differences are (1) the identities and locations of the electrophilic cationic residue that interacts with the substrate carboxylate group (Lys 283 in ManD and Lys 164 in MR), and (2) the participation of Arg 147 with Tyr 159 in a hydrogen-bonded dyad that functions as the base in the ManD-catalyzed reaction (Lys 166 is the (*S*)-specific base in the MR-catalyzed reaction). Also, although both active sites contain a His at the end of the seventh β-strand, in ManD His 312 does not directly participate in catalysis and is not hydrogen-bonded to an Asp at the end of the sixth β-strand; in MR His 297 is hydrogen-bonded to Asp 270 and functions as the (*R*)-specific base. The essential Mg^{2+} ions are shown as spheres.

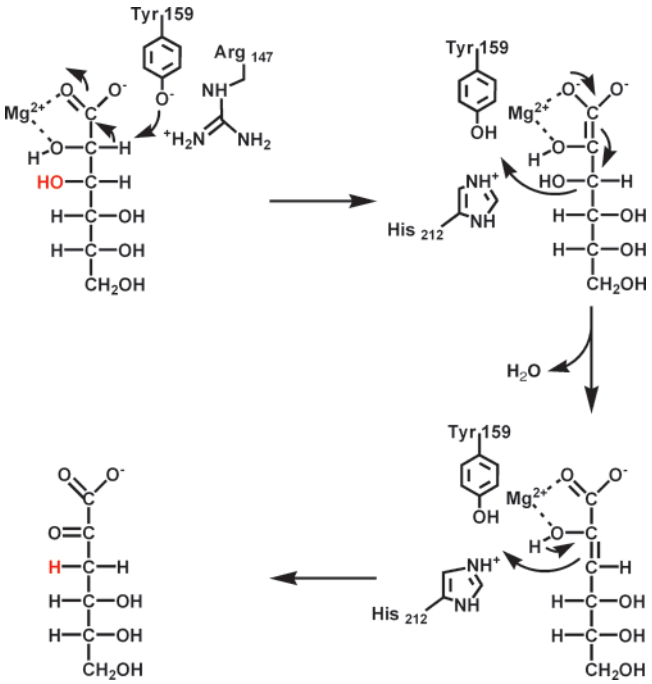


FIGURE 9: Proposed mechanism of the ManD-catalyzed reaction. Tyr 159, hydrogen-bonded to Arg 147, is positioned to function as the base that abstracts the 2-proton to generate the stabilized enediolate intermediate. The vinylogous syn-elimination of the 3-OH followed by ketonization with retention of configuration to generate the 2-keto-3-deoxy-D-gluconate product may be catalyzed by Tyr 159 and/or His 212.

Comparison with Acid Sugar Dehydratases in the MR Subgroup. FucD and TarD catalyze the anti-dehydration of their substrates, using the Lys at end of the second β-strand as the base that abstracts the 2-proton and the His at the end of the seventh β-strand that facilitates departure of the 3-OH group. TalrD/GalrD, a promiscuous enzyme, catalyzes the anti-dehydration of L-talarate and the syn-dehydration of galactarate: the dehydration of L-talarate is initiated by the Lys at the end of the second β-strand, and the dehydration of galactarate is initiated by the His at the end of the seventh β-strand; for both substrates, the His at the end of the seventh β-strand facilitates departure of the 3-OH group (*24*). Interestingly, FucD catalyzes replacement of the 3-OH group with solvent deuterium with inversion of configuration (consistent with the Lys at the end of the second β-strand also participating in this reaction) (*22*), and TalrD/GalrD catalyzes replacement of the 3-OH group with solvent deuterium with retention of configuration (consistent with the His at the end of the seventh β-strand also participating in this reaction). Although these active sites use different structural strategies for catalyzing the ketonization of the enol intermediate derived from dehydration, formation of the 2-keto dehydration product is enzyme-catalyzed.

TarD catalyzes the anti-elimination of the 3-OH group to generate an enol "intermediate" analogous to those produced by FucD and TalrD/GalrD, also using the Lys at the end of the second β-strand as the base and the His at the end of the seventh β-strand as the acid (*23*). However, this enol is released as product, as judged by both direct observation of the enol and stereorandom incorporation of solvent deuterium in the oxaloacetate product.

Thus, although the structural strategy for stabilization of the essential, unstable enolate anion intermediate derived by abstraction of the 2-proton is conserved in ManD and the acid sugar dehydratases of the MR subgroup, the structural strategies for generating that intermediate and directing it to the product that is released from the enzyme are not conserved. These differences illustrate the changes in mechanism that are possible in the process of divergent evolution of similar enzymatic functions in the enolase superfamily.

Conclusions. As functions are assigned to previously uncharacterized members of the enolase superfamily discovered in genome projects, we are recognizing that divergent evolution of function can involve structural and functional alterations that cannot be predicted based on previously determined paradigms. Indeed, the New York Structural Genomics Consortium supported by the Protein Structure Initiative has solved and deposited in the Protein Data Bank structures for several uncharacterized and, as yet, functionally unassigned members of the enolase superfamily that reveal additional changes in structural details, although the bidomain structure, the N-terminal $\alpha+\beta$ capping domain and the C-terminal $(\beta/\alpha)_7\beta$-barrel domain, as well as the positions of the carboxylate ligands for the essential Mg^{2+} are conserved.

Given the structural variations that have been observed in enolase superfamily, we recognize that the only conserved residues are ligands for the essential Mg^{2+} at the ends of the third, fourth, and fifth β-strands of the barrel domain (Table 1). Given the exceptional values for the pK_as of the protons that are abstracted ($\sim$29 for mandelate anion and $\sim$32 for acid sugars) (*38*, *39*), Nature conserved the strategy for stabilizing the enolate anion intermediates, although the identities and spatial locations of the bases that catalyze proton abstraction to generate the intermediate are not conserved.

ACKNOWLEDGMENT

We acknowledge Cheri Millikin for her assistance with early aspects of this work and Dr. Shoshana Brown for her assistance with the sequence alignment shown in Figure 3.

REFERENCES

1. Pegg, S. C., Brown, S. D., Ojha, S., Seffernick, J., Meng, E. C., Morris, J. H., Chang, P. J., Huang, C. C., Ferrin, T. E., and Babbitt, P. C. (2006) Leveraging enzyme structure-function relationships for functional inference and experimental design: the structure-function linkage database, *Biochemistry 45*, 2545−55.
2. Lamble, H. J., Milburn, C. C., Taylor, G. L., Hough, D. W., and Danson, M. J. (2004) Gluconate dehydratase from the promiscuous Entner-Doudoroff pathway in Sulfolobus solfataricus, *FEBS Lett. 576*, 133−6.
3. Ahmed, H., Ettema, T. J., Tjaden, B., Geerling, A. C., van der Oost, J., and Siebers, B. (2005) The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic archaea: a re-evaluation, *Biochem. J. 390*, 529−40.
4. Wieczorek, S. W., Kalivoda, K. A., Clifton, J. G., Ringe, D., Petsko, G. A., and Gerlt, J. A. (1999) Evolution of Enzymatic Activities in the Enolase Superfamily: Identification of a "New" General Acid Catalyst in the Active Site of D-Galactonate Dehydratase from Escherichia coli, *J. Am. Chem. Soc. 121*, 4540−4541.
5. Babbitt, P. C., Mrachko, G. T., Hasson, M. S., Huisman, G. W., Kolter, R., Ringe, D., Petsko, G. A., Kenyon, G. L., and Gerlt, J. A. (1995) A functionally diverse enzyme superfamily that abstracts the alpha protons of carboxylic acids, *Science 267*, 1159−61.
6. Gerlt, J. A., and Babbitt, P. C. (2001) Divergent Evolution of Enzymatic Function: Mechanistically Diverse Superfamilies and Functionally Distinct Suprafamilies, *Annu. Rev. Biochem. 70*, 209−246.
7. Gerlt, J. A., Babbitt, P. C., and Rayment, I. (2005) Divergent evolution in the enolase superfamily: the interplay of mechanism and specificity, *Arch. Biochem. Biophys. 433*, 59−70.
8. Glasner, M. E., Gerlt, J. A., and Babbitt, P. C. (2006) Evolution of enzyme superfamilies, *Curr. Opin. Chem. Biol. 10*, 492−7.
9. Babbitt, P. C., and Gerlt, J. A. (1997) Understanding enzyme superfamilies. Chemistry as the fundamental determinant in the evolution of new catalytic activities, *J. Biol. Chem. 272*, 30591−4.
10. Larsen, T. M., Wedekind, J. E., Rayment, I., and Reed, G. H. (1996) A carboxylate oxygen of the substrate bridges the magnesium ions at the active site of enolase: structure of the yeast enzyme complexed with the equilibrium mixture of 2-phosphoglycerate and phosphoenolpyruvate at 1.8 A resolution, *Biochemistry 35*, 4349−58.
11. Reed, G. H., Poyner, R. R., Larsen, T. M., Wedekind, J. E., and Rayment, I. (1996) Structural and mechanistic studies of enolase, *Curr. Opin. Struct. Biol. 6*, 736−43.
12. Levy, C. W., Buckley, P. A., Sedelnikova, S., Kato, Y., Asano, Y., Rice, D. W., and Baker, P. J. (2002) Insights into enzyme evolution revealed by the structure of methylaspartate ammonia lyase, *Structure (Camb) 10*, 105−13.
13. Asuncion, M., Blankenfeldt, W., Barlow, J. N., Gani, D., and Naismith, J. H. (2002) The structure of 3-methylaspartase from Clostridium tetanomorphum functions via the common enolase chemical step, *J. Biol. Chem. 277*, 8306−11.
14. Hasson, M. S., Schlichting, I., Moulai, J., Taylor, K., Barrett, W., Kenyon, G. L., Babbitt, P. C., Gerlt, J. A., Petsko, G. A., and Ringe, D. (1998) Evolution of an enzyme active site: the structure of a new crystal form of muconate lactonizing enzyme compared with mandelate racemase and enolase, *Proc. Natl. Acad. Sci. U.S.A. 95*, 10396−401.
15. Thompson, T. B., Garrett, J. B., Taylor, E. A., Meganathan, R., Gerlt, J. A., and Rayment, I. (2000) Evolution of enzymatic activity in the enolase superfamily: structure of *o*-succinylbenzoate synthase from Escherichia coli in complex with Mg^{2+} and *o*-succinylbenzoate, *Biochemistry 39*, 10662−76.
16. Schmidt, D. M., Hubbard, B. K., and Gerlt, J. A. (2001) Evolution of Enzymatic Activities in the Enolase Superfamily: Functional Assignment of Unknown Proteins in Bacillus subtilis and Escherichia coli as L-Ala-D/L-Glu Epimerases, *Biochemistry 40*, 15707−15715.
17. Gulick, A. M., Schmidt, D. M., Gerlt, J. A., and Rayment, I. (2001) Evolution of Enzymatic Activities in the Enolase Superfamily: Crystal Structures of the L-Ala-D/L-Glu Epimerases from Escherichia coli and Bacillus subtilis, *Biochemistry 40*, 15716−15724.
18. Taylor, Ringia, E. A., Garrett, J. B., Thoden, J. B., Holden, H. M., Rayment, I., and Gerlt, J. A. (2004) Evolution of enzymatic activity in the enolase superfamily: functional studies of the promiscuous *o*-succinylbenzoate synthase from Amycolatopsis, *Biochemistry 43*, 224−9.
19. Thoden, J. B., Taylor Ringia, E. A., Garrett, J. B., Gerlt, J. A., Holden, H. M., and Rayment, I. (2004) Evolution of Enzymatic Activity in the Enolase Superfamily: Structural Studies of the Promiscuous *o*-Succinylbenzoate Synthase from Amycolatopsis, *Biochemistry 43*, 5716−27.
20. Sakai, A., Xiang, D. F., Xu, C., Song, L., Yew, W. S., Raushel, F. M., and Gerlt, J. A. (2006) Evolution of enzymatic activities in the enolase superfamily: *N*-succinylamino acid racemase and a new pathway for the irreversible conversion of D- to L-amino acids, *Biochemistry 45*, 4455−62.
21. Neidhart, D. J., Howell, P. L., Petsko, G. A., Powers, V. M., Li, R. S., Kenyon, G. L., and Gerlt, J. A. (1991) Mechanism of the reaction catalyzed by mandelate racemase. 2. Crystal structure of mandelate racemase at 2.5-Å resolution: identification of the active site and possible catalytic residues, *Biochemistry 30*, 9264−73.
22. Yew, W. S., Fedorov, A. A., Fedorov, E. V., Rakus, J. F., Pierce, R. W., Almo, S. C., and Gerlt, J. A. (2006) Evolution of enzymatic activities in the enolase superfamily: L-fuconate dehydratase from Xanthomonas campestris, *Biochemistry 45*, 14582−97.
23. Yew, W. S., Fedorov, A. A., Fedorov, E. V., Wood, B. M., Almo, S. C., and Gerlt, J. A. (2006) Evolution of enzymatic activities in the enolase superfamily: D-tartrate dehydratase from Bradyrhizobium japonicum, *Biochemistry 45*, 14598−608.
24. Yew, W. S., Fedorov, A. A., Fedorov, E. V., Almo, S. C., and Gerlt, J. A. (2007) Evolution of Enzymatic Activities in the Enolase Superfamily: L-Talarate/Galactarate Dehydratase from Salmonella typhimurium LT2, *Biochemistry 46*, 9564−77.
25. Palmer, D. R., and Gerlt, J. A. (1996) Evolution of Enzymatic Activities: Multiple Pathways for Generating and Partitioning a Common Enolic Intermediate by Glucarate Dehydratase from Pseudomonas putida, *J. Am. Chem. Soc. 118*, 10323−10324.
26. Gulick, A. M., Palmer, D. R., Babbitt, P. C., Gerlt, J. A., and Rayment, I. (1998) Evolution of enzymatic activities in the enolase superfamily: crystal structure of (D)-glucarate dehydratase from Pseudomonas putida, *Biochemistry 37*, 14358−68.
27. Song, L., Kalyanaraman, C., Fedorov, A. A., Fedorov, E. V., Glasner, M. E., Brown, S., Imker, H. J., Babbitt, P. C., Almo, S. C., Jacobson, M. P., and Gerlt, J. A. (2007) Prediction and assignment of function for a divergent N-succinyl amino acid racemase. *Nat. Chem. Biol. 3*, 486−91.
28. Kenyon, G. L., Gerlt, J. A., Petsko, G. A., and Kozarich, J. W. (1995) Mandelate Racemase: Structure-function studies of a pseudosymmetric enzyme, *Acc. Chem. Res. 28*, 178−186.
29. Williams, L., Nguyen, T., Li, Y., Porter, T. N., and Raushel, F. M. (2006) Uronate isomerase: a nonhydrolytic member of the amidohydrolase superfamily with an ambivalent requirement for a divalent metal ion, *Biochemistry 45*, 7453−62.
30. Gulick, A. M., Hubbard, B. K., Gerlt, J. A., and Rayment, I. (2000) Evolution of enzymatic activities in the enolase superfamily: crystallographic and mutagenesis studies of the reaction catalyzed by D- glucarate dehydratase from Escherichia coli, *Biochemistry 39*, 4590−602.
31. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, in *Methods Enzymol.* (Carter, C. W. J., Sweet, R. M., Abelson, J. N., and Simon, M. I., Eds.) pp 307−326, Academic Press, New York.
32. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution, *Acta Crystallogr. D: Biol. Crystallogr. 55*, 849−61.
33. Terwilliger, T. C. (2000) Maximum-likelihood density modification, *Acta Crystallogr. D: Biol. Crystallogr. 56*, 965−72.
34. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol. 6*, 458−63.
35. Jones, A. T. (1985) Interactive computer graphics: FRODO, *Methods Enzymol. 115*, 157−71.
36. 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) Crystallography & NMR system: A new

software suite for macromolecular structure determination, *Acta Crystallogr. D54*.
37. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions, *Acta Crystallogr D. Biol. Crystallogr 61*, 458−64.
38. Gerlt, J. A., Kozarich, J. W., Kenyon, G. L., and Gassman, P. G. (1991) Electrophilic Catalysis Can Explain the Unexpected Acidity of Carbon Acids in Enzyme-Catalyzed Reactions, *J. Am. Chem. Soc. 113*, 9667−69.
39. Richard, J. P., and Amyes, T. L. (2001) Proton transfer at carbon, *Curr. Opin. Chem. Biol. 5*, 626−33.

BI701703W