Evolution of Enzymatic Activities in the Enolase Superfamily: L-Rhamnonate Dehydratase[†,‡]

John F. Rakus,[§] Alexander A. Fedorov,[||,⊥] Elena V. Fedorov,[||] Margaret E. Glasner,[#] Brian K. Hubbard,[§] Joseph D. Delli,[§] Patricia C. Babbitt,[#] Steven C. Almo,[*,||,⊥] and John A. Gerlt[*,§]

Departments of Biochemistry and Chemistry, University of Illinois at Urbana−Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, Department of Biochemistry and New York SGX Research Center for Structural Genomics, Albert Einstein College of Medicine, Bronx, New York 10461, and Department of Biopharmaceutical Sciences, School of Pharmacy and California Institute for Quantitative Biomedical Research, University of California, 1700 Fourth Street, San Francisco, California 94158

Received May 15, 2008; Revised Manuscript Received June 18, 2008

ABSTRACT: The L-rhamnonate dehydratase (RhamD) function was assigned to a previously uncharacterized family in the mechanistically diverse enolase superfamily that is encoded by the genome of *Escherichia coli* K-12. We screened a library of acid sugars to discover that the enzyme displays a promiscuous substrate specificity: L-rhamnonate (6-deoxy-L-mannonate) has the "best" kinetic constants, with L-mannonate, L-lyxonate, and D-gulonate dehydrated less efficiently. Crystal structures of the RhamDs from both *E. coli* K-12 and *Salmonella typhimurium* LT2 (95% sequence identity) were obtained in the presence of Mg^{2+}; the structure of the RhamD from *S. typhimurium* was also obtained in the presence of 3-deoxy-L-rhamnonate (obtained by reduction of the product with $NaBH_4$). Like other members of the enolase superfamily, RhamD contains an N-terminal $\alpha + \beta$ capping domain and a C-terminal $(\beta/\alpha)_7\beta$-barrel (modified TIM-barrel) catalytic domain with the active site located at the interface between the two domains. In contrast to other members, the specificity-determining "20s loop" in the capping domain is extended in length and the "50s loop" is truncated. The ligands for the Mg^{2+} are Asp 226, Glu 252 and Glu 280 located at the ends of the third, fourth and fifth β-strands, respectively. The active site of RhamD contains a His 329-Asp 302 dyad at the ends of the seventh and sixth β-strands, respectively, with His 329 positioned to function as the general base responsible for abstraction of the C2 proton of L-rhamnonate to form a Mg^{2+}-stabilized enediolate intermediate. However, the active site does not contain other acid/base catalysts that have been implicated in the reactions catalyzed by other members of the MR subgroup of the enolase superfamily. Based on the structure of the liganded complex, His 329 also is expected to function as the general acid that both facilitates departure of the 3-OH group in a *syn*-dehydration reaction and delivers a proton to carbon-3 to replace the 3-OH group with retention of configuration.

The members of the enolase superfamily catalyze mechanistically diverse reactions that are initiated by base-assisted abstraction of the α-proton of a carboxylate anion substrate to form an enediolate intermediate; the intermediate must be stabilized by an essential Mg^{2+} ion to be kinetically competent (*1*−*3*). The members share a conserved tertiary structure with a two-domain architecture, in which three carboxylate ligands for the Mg^{2+} ion as well as the acid/base catalysts are located at the C-terminal ends of the β-strands in a $(\beta/\alpha)_7\beta$-barrel [modified $(\beta/\alpha)_8$- or TIM-barrel] domain and the specificity-determining residues are located in an N-terminal $\alpha + \beta$ capping domain.

The pseudosymmetric structure of the barrel domain provides a mechanistically versatile scaffold for the catalytic residues (*4*). Functional groups are positioned on either face of the enediolate intermediate to allow mechanistic variation that is characterized by an array of stereochemical characteristics. The known reactions involve β-elimination of either hydroxide or ammonia leaving groups, 1,1-proton transfer in racemization and epimerization, and intramolecular addition/elimination (cycloisomerization).

Almost all members of the superfamily contain three carboxylate ligands for the essential Mg^{2+} ion separately located at the C-terminal ends of the third, fourth and fifth β-strands of the barrel-domain. The identities and positions

[†] This research was supported by GM-52594, P01 GM-71790, U54 GM-74945, and a Ruth L. Kirschstein National Research Service Award 5 T32 GM070421 from the National Institutes of Health.

[‡] The X-ray coordinates and structure factors for RhamD from *Salmonella typhimurium* LT2 liganded with Mg^{2+}, RhamD from *S. typhimurium* LT2 liganded with Mg^2 and 3-deoxy-L-rhamnonate, and RhamD from *Escherichia coli* in the absence of ligands have been deposited in the Protein Data Bank (PDB accession codes, 3BOX, 3CXO, and 2I5Q, respectively).

* To whom correspondence should be addressed. J.A.G.: Department of Biochemistry, University of Illinois, 600 S. Mathews Avenue, Urbana, IL 61801. Phone: (217) 244-7414. Fax: (217) 244-6538. E-mail: j-gerlt@uiuc.edu. S.C.A.: Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461. Phone: (718) 430-2746. Fax: (718) 430-8565. E-mail: almo@aecom.yu.edu.

[§] University of Illinois at Urbana−Champaign.

[||] Department of Biochemistry, Albert Einstein College of Medicine.

[⊥] New York SGX Research Center for Structural Genomics, Albert Einstein College of Medicine.

[#] University of California.

10.1021/bi800914r CCC: $40.75
Published on Web 08/29/2008

of the acid/base catalysts allow the superfamily to be divided into subgroups that differ in their structural strategies for catalysis.

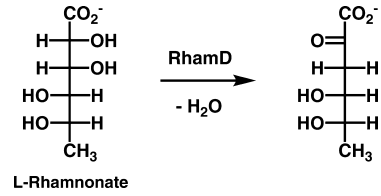
Two subgroups share a His-Asp dyad at the ends of the seventh and sixth β-strands, respectively (*5*). In the *mandelate racemase (MR[1]) subgroup*, one carboxylate oxygen and the α-OH group of the substrate are ligands for the essential Mg^{2+}. In addition to the MR function (1,1-proton transfer), several acid sugar dehydratases (β-elimination) have been identified in the MR subgroup, including D-arabinonate dehydratase (AraD (*6*)), L-fuconate dehydratase (FucD (*7*)), D-galactonate dehydratase (GalD (*8*)), D-gluconate dehydratase (GlcD (*9, 10*)), D-tartrate dehydratase (TarD (*11*)), and L-talarate/galactarate dehydratase (TalrD/GalrD (*12*)); structures are available for MR (*13*), GalD (*8*), TarD (*11*), FucD (*7*), and TalrD/GalrD (*12*). Acid/base catalysts located at various positions in the barrel domain are responsible for the differing stereochemical courses of the dehydration reactions (e.g., abstraction of a 2*R*- or 2*S*-proton; *syn*- or *anti*-elimination of the 3-OH group).

Although the *D-glucarate dehydratase (GlucD) subgroup* shares the His-Asp dyad with the MR subgroup, the ligand for the essential Mg^{2+} at the end of the fifth β-strand is an Asn instead of a Glu, with the consequence that the carboxylate group of the substrate is a bidentate ligand for the essential Mg^{2+} (*14*). On the basis of the structure of a substrate-analogue liganded complex, the resulting differing disposition of the α-proton and 3-OH group *vis-à-vis* the acid/base catalysts results in a different geometric solution to dehydration, with the His-Asp dyad catalyzing not only abstraction of the 2-proton from D-glucarate but also the vinylogous elimination of the 3-OH group and final delivery of a solvent-derived hydrogen to carbon-3.

Seventeen different reactions have been assigned to various members of the enolase superfamily (*5*); however, at least 50% of the members have unknown or uncertain functions (*15*). Some organisms contain only one member of the superfamily, i.e., the ubiquitous enolase, but others contains several, with at least one usually having an unknown or uncertain function. As an "extreme" example, the genome of *Solibacter usitatus* Ellin6075 encodes 25 members of the superfamily, but only two can be assigned a function based on sequence homology. In contrast, the genome of *Escherichia coli* K-12 encodes eight members, six of which have been assigned a function based on enzymatic assays and/or sequence homology: enolase (GI:16130868), *o*-succinylbenzoate synthase (GI:16130196 (*16*)), L-Ala-D/L-Glu epimerase (GI:90111249 (*17, 18*)), GalD (GI:49176390 (*8*)), GlucD (GI:16130695 (*14, 19*)), and ManD (GI:16129539 (*5*)).

We now describe structural characterization of a seventh member encoded by the *E. coli* K-12 genome (GI:16130182, designated YfaW and annotated as a "putative racemase") and its functional assignment as L-rhamnonate dehydratase (RhamD; Scheme 1).

Scheme 1



RhamD is a member of the MR subgroup, with three carboxylate ligands for the essential Mg^{2+} (Asp 226, Glu 252, and Glu 280) and a His-Asp dyad (His 329 and Asp 302). Based on the structural and mechanistic data, we conclude that RhamD catalyzes a *syn*-dehydration reaction. The His-Asp dyad is appropriately located to catalyze all three partial reactions: (1) abstraction of the proton from carbon-2 to generate an enolate anion intermediate; (2) acid catalysis of vinylogous departure of the 3-OH group; and (3) protonation of carbon-3 to replace the 3-OH group with a solvent-derived hydrogen with retention of configuration. The structures of the specificity determining loops in the capping domain differ from those previously observed in members of the MR subgroup, thereby establishing a new structural strategy for evolution of function in the enolase superfamily.

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.

3-Deoxy-L-Rhamnonate. L-Rhamnonate (800 mg, 14.6 mmol) was incubated with RhamD (220 μM) in 25 mL of 50 mM Tris-HCl, pH 7.9, containing 5 mM $MgCl_2$ at room temperature for 72 h. Completion of the reaction was verified by ^{1}H NMR. The enzyme was removed by ultrafiltration, sodium borohydride (800 mg, 4.3 mmol) was added, and the reaction was stirred at room temperature for 3 h. The reaction was terminated by adjusting the pH to 7 to afford a mixture of 3-deoxy-L-rhamnonate and its 2-epimer. Attempts to separate the compounds by chromatography were unsuccessful, so the mixture was used for cocrystallization experiments.

Cloning, Expression, and Purification of YfaW (GI: 16130182) from E. coli K-12. The gene encoding YfaW (GI: 16130182) was PCR-amplified from *E. coli* MG1655 genomic DNA using *Pfu* DNA polymerase (Stratagene). A PTC-200 gradient thermal cycler (MJ Research) was used with the following parameters: 40 cycles of denaturation at 95 °C for 1 min and annealing/extension at 60 °C for 1 min followed by a final extension at 72 °C extension for two min. The blunted-ended products were isolated with a Qiagen PCR purification kit, 5′-phosphorylated with polynucleotide kinase, and ligated into either pET15b (encoding an N-terminal His_6-tag) or a modified pET15b (encoding an N-terminal His_{10}-tag) that had been linearized with *Stu*I. The proteins were expressed in *E. coli* strain BL21(DE3). Transformed cells were grown overnight at 37 °C in LB supplemented with 100 μg/mL ampicillin and harvested by centrifugation (5000 rpm at 4 °C); no induction with IPTG

[1] Abbreviations: AraD, D-arabinonate dehydratase; FucD, L-fuconate dehydratase; GalD, D-galactonate dehydratase; GlcD, D-gluconate 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-methylaspartate ammonia lyase; ManD, D-mannonate dehydratase; MLE, muconate lactonizing enzyme; MR, mandelate racemase; NYSGXRC, New York SGX Research Center for Structural Genomics; PK, pyruvate kinase; RhamD, L-rhamnonate dehydratase; ecRhamD, RhamD from *Escherichia coli* K-12; stRhamD, RhamD from *Salmonella typhimurium* LT2; TarD, D-tartrate dehydratase; TalrD/GalrD, L-talarate/galactarate dehydratase; TIM, triose phosphate isomerase.

was performed. Each His-tagged protein was purified with a Chelating Sepharose Fast Flow (Pharmacia) column (16 cm × 40 cm). The cell lysate was applied to the column in binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9), washed with wash buffer (60 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9), and eluted with 50% binding buffer, 50% strip buffer (100 mM EDTA, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9).

Expression and Purification of YfaU (GI:16130180) from E. coli K-12. The gene encoding YfaU (GI:16130180) from *E. coli* K-12 was acquired from the ASKA Collection in the pCA24N vector that contains a C-terminal His_6-tag. The protein was expressed in *E. coli* BL21(DE3) cells grown overnight at 37 °C in LB supplemented with 34 μg/mL chloramphenicol. The cells were harvested by centrifugation and lysed by sonication. The lysate was purified on a Chelating Sepharose Fast Flow column as described for YfaW.

Site-Directed Mutagenesis. The H281N, H282N and H329N mutants of the gene encoding YfaW from *E. coli* were constructed as follows. Two PCR reactions were performed for each mutation: (1) the 5′-sense primer encoding the mutant was paired with an antisense-strand primer that binds downstream of the termination codon; and (2) the 3′-antisense primer encoding the mutant was paired with a sense-strand primer that binds upstream of the initiation codon. The product obtained for the first reaction was digested with *Xho*I and 5′-phosphorylated. The second PCR (with the T7 promoter primer) was digested with *Xba*I and 5′-phosphorylated. The two digested and phosphorylated pieces were ligated into a pET vector that has been digested with *Xba*I and *Xho*I, and dephosphorylated.

The H33N mutant of the gene encoding YfaW from *E. coli* was constructed by the overlap extension method. The megaprimers were constructed in reactions (50 μL) that contained 1 ng of the plasmid encoding wild type RhamD in the pET15b vector, 5 μL of 10x PCR buffer, 4 mM $MgCl_2$, 2 mM dNTPs, 40 pmol of each primer, 1 unit of *Taq* DNA polymerase (Invitrogen) and 0.5 unit of cloned *Pfu* polymerase (Stratagene). The 5′-megaprimer was constructed with T7 promoter primer and a second 3′-primer that encodes the mutant; the 3′-megaprimer was constructed the T7 terminator primer and a second 5′-primer that encodes the mutant. The parameters for the PCR reactions were 95 °C for 4 min followed by 26 cycles of 95 °C for 45 s, 55 °C for 45 s, and 72 °C for 135 s followed finally by 7 min at 72 °C. After purification by 1% agarose gel electrophoresis, the final PCR reaction (50 μL) using the same program contained 5 μL 10x PCR buffer, 4 mM $MgCl_2$, 2 mM dNTPs, 40 pmol of T7 promoter and T7 terminator primers, 200 pmol of each of the megaprimers, 1 unit of *Taq* DNA polymerase (Invitrogen) and 0.5 unit of cloned *Pfu* polymerase (Stratagene).

Screen for Dehydration Activity. YfaW was screened for dehydration activity using a library of mono- and diacid sugars as described previously (*7*).

CycleNOE Spectrum of KDR. To identify the absolute configuration of the C-3 protons, the C-6 methyl group was irradiated in a ^1H cycleNOE experiment. The proton signal which was irradiated in the energy transfer was the proton on the same face of the cyclic KDR product as the C-6 methyl group.

Assay for RhamD Activity. YfaW was assayed at 25 °C either (1) discontinuously by quenching 100 μL of a reaction (containing 5 mM $MgCl_2$, 50 mM Tris-HCl, pH 7.9, and 250 nM RhamD) with 400 μL of 1% semicarbazide/1% sodium acetate; or (2) continuously by a coupled-enzyme spectrophotometric assay (1 mL) containing 50 mM potassium HEPES, pH 7.5, 5 mM $MgCl_2$, 0.16 mM NADH, 10 U of LDH, and 400 nM YfaU and monitoring NADH consumption at 340 nm.

Assay for 2-Keto-3-deoxy-L-rhamnonate Aldolase Activity. A stock solution (100 mM) of 2-keto-3-deoxy-L-rhamnonate (the product of the RhamD-catalyzed reaction, KDR) was obtained by incubating a solution (4 mL) containing 100 mM L-rhamnonate, 50 mM Tris-HCl, pH 8.0, and 5 mM $MgCl_2$ with RhamD (1 μM) overnight at room temperature. The enzyme was removed by filtration using an Amicon Micropure-EZ centrifugal filtration device at 3000 rpm for 90 min; completion of the reaction was confirmed by ^1H NMR. Samples of 2-keto-3-deoxy-L-lyxonate, and 2-keto-3-deoxy-L-mannonate were analogously prepared.

The aldolase activity was quantitated at 25 °C with a coupled-enzyme spectrophotometric assay. The reactions (1 mL) contained 5 mM $MgCl_2$, 50 mM potassium Hepes, pH 7.5, 0.16 mM NADH, 10 U of lactate dehydrogenase (LDH), and varying concentrations of the 2-keto-3-deoxy acid sugar substrates. Consumption of NADH was quantitated at 340 nm.

Assay for 2-Keto-hept-3-ene-1,7-dioate Hydratase Activity (20). One hundred microliters of an ethanolic solution of 2-hydroxy-2,4-diene-1,7-dioate (4 mg/mL) was added to 900 μL of 20 mM potassium phosphate, pH 7.3, containing 5 mM $MgCl_2$; 4-oxalocrotonate tautomerase was added to generate 2-keto-hept-3-ene-1,7-dioate. A 20 μL aliquot of this solution was added to a cuvette containing 1 mL of 20 mM potassium phosphate, pH 7.3, and 5 mM $MgCl_2$; hydration activity was quantitated at 25 °C by the decrease in absorbance at 232 nm ($\varepsilon = 8250\ M^{-1}\ cm^{-1}$).

Crystallization and Data Collection. Three different crystal forms (Table 1) were grown by the hanging drop method at room temperature: (1) RhamD from *S. typhimurium* (stRhamD) and Mg^{2+}, (2) stRhamD, Mg^{2+}, and 3-deoxy-L-rhamnonate, and (3) RhamD from *E. coli* K-12 (ecRhamD) in the absence of ligands.

For stRhamD and Mg^{2+}, the protein solution contained stRhamD (9.7 mg/mL) in 10 mM HEPES, pH 7.5, containing 150 mM NaCl, 10 mM L-methionine, 5 mM DTT, 10% glycerol, and 5 mM $MgCl_2$; the precipitant contained 2.4 M sodium malonate, pH 7.0, and 5 mM $MgCl_2$. Crystals appeared in 3−4 days and exhibited diffraction consistent with the space group *F*432, with two molecules of stRhamD per asymmetric unit.

For stRhamD, Mg^{2+}, and 3-deoxy-L-rhamnonate, the protein solution contained stRhamD (42 mg/mL) in 10 mM HEPES, pH 7.5, containing 150 mm NaCl, 10 mM L-methionine, 5 mM DTT, 10% glycerol, 5 mM $MgCl_2$, and 40 mM 3-deoxy-L-rhamnonate; the precipitant contained 60% Tacsimate, pH 7.0, and 5 mM $MgCl_2$. Crystals appeared in 3−4 days and exhibited diffraction consistent with the space group *F*432, with two molecules of stRhamD per asymmetric unit.

For ecRhamD, the protein solution contained ecRhamD (42 mg/mL) in 10 mM HEPES, pH 7.5, containing 150 mM

Table 1: Data Collection and Refinement Statistics

	stRhamD·Mg^{2+}	stRhamD·Mg^{2+}·3-d-L-Rham	ecRhamD
	Data Collection		
wavelength (Å)	0.979	0.979	0.979
space group	*F*432	*F*432	*I*4
mol in au	2	2	2
unit cell parameters			
a (Å)	271.96	272.83	124.82
c (Å)			103.76
resolution (Å)[a]	30−1.8 (1.86−1.8)	25.0−2.0 (2.07−2.0)	2.1 (2.18−2.10)
unique reflections	77763	57744	44477
completeness (%)[a]	97.7 (85.8)	98.1 (96.6)	95.8 (88.1)
R_{merge}[a]	0.082 (0.355)	0.071 (0.285)	0.065 (0.343)
average I/σ[a]	36.9 (9.6)	30.8 (7.1)	15.8 (5.4)
	Refinement		
resolution (Å)	25−1.8	25−2.0	25−2.1
R_{cryst}	0.201	0.221	0.229
R_{free}	0.224	0.234	0.258
rmsd, bonds (Å)	0.006	0.007	0.006
rmsd, angles (deg)	1.35	1.39	1.43
no. of atoms			
protein	6223	6225	5930
water	459	436	110
Mg^{2+}	2	2	
bound ligand		2	
PDB entry	3BOX	3CXO	2I5Q

[a] Numbers in parentheses indicate values for the highest resolution shell.

NaCl, 10 mM L-methionine, and 10% glycerol; the precipitant contained 1.2 M ammonium sulfate, 100 mM Tris-HCl, pH 8.5, and 0.2 M lithium sulfate. Crystals appeared in 2 days and exhibited diffraction consistent with space group *I*4, with two molecules of ecRhamD per asymmetric unit.

Prior to data collection, the crystals were transferred to a cryoprotectant solution composed of their mother liquid and 20% glycerol and flash-cooled in a nitrogen stream. X-ray diffraction data sets for stRhamD with Mg^{2+} (Table 1, column 1); the complex of stRhamD with Mg^{2+} and 3-deoxy-L-rhamnonate (column 2); and the ecRhamD (column 3) were collected at the NSLS X4A beamline (Brookhaven National Laboratory) on an ADSC CCD detector to 1.8, 2.0, and 2.1 Å resolution, respectively. Diffraction intensities were integrated and scaled with programs DENZO and SCALEPACK (*21*). The data collection statistics are given in Table 1.

Structure Determination and Refinement. The structure of ecRhamD was solved by molecular replacement with the fully automated molecular replacement pipeline BALBES (*22*), using as input only the diffraction and sequence data; the partially refined structure was obtained without any manual intervention. Subsequent iterative cycles of manual rebuilding with TOM (*23*), refinement with CNS (*24*), and automatic rebuilding with ARP (*25*) resulted in a model with R_{cryst} and R_{free} 0.229 and 0.258, respectively. The final structure contained 5930 protein atoms and 110 water molecules for one dimer in the asymmetric unit. One nonglycine residue, Thr 284, was located in the disallowed region of the Ramachandran plot for both monomers; this residue had well-defined density and is located at the intermolecular interface. Residues 24−38 and 57−60 have no density in both monomers and were not included in the final model.

The structure of stRhamD crystallized with Mg^{2+} was automatically solved and partially refined with BALBES using diffraction and sequence data. Subsequent iterative cycles of manual rebuilding with TOM, refinement with CNS, and automatic rebuilding with ARP were performed. The model was refined at 1.8 Å with an R_{cryst} of 0.201 and an R_{free} of 0.224. The final structure contained residues 4−405 and well-defined Mg^{2+} ions in both monomers of the dimer. All residues in 20s and 50s loops in both monomers were well-defined. The Mg^{2+} ions are coordinated by side chains of Asp 226, Glu 252, Glu 280, and three water molecules.

The protein part of the structure of stRhamD crystallized with Mg^{2+} was used as starting point for refinement of stRhamD crystallized with Mg^{2+} and 3-deoxy-L-rhamnonate. Iterative cycles of manual rebuilding with TOM, refinement with CNS, and automatic rebuilding with ARP with subsequent inclusion of water molecules were performed. Mg^{2+} ions were clearly defined in both monomers and have octahedral coordination in each. The active site of polypeptide A had well-defined density for 3-deoxy-L-rhamnonate. The active site of polypeptide B contained unexpected elongated electron density, resembling that expected for 2,4-dihydroxyoctanoate although this molecule was not present in the crystallization solution; this electron density might be explained by statistical superposition of 3-deoxy-L-rhamnonate and L-methionine that also was present in the protein solution.

Final crystallographic refinement statistics are provided in Table 1.

RESULTS AND DISCUSSION

As summarized in the introductory comments, six of the eight members of the enolase superfamily encoded by the *E. coli* K-12 genome have known functions. This manuscript describes the functional assignment and structural characterization of the seventh, YfaW (GI:16130182).

Predictions from Sequence Alignments. The sequence of YfaW contains the conserved His-Asp dyad motif found in members of the MR and GlucD subgroups. An alignment

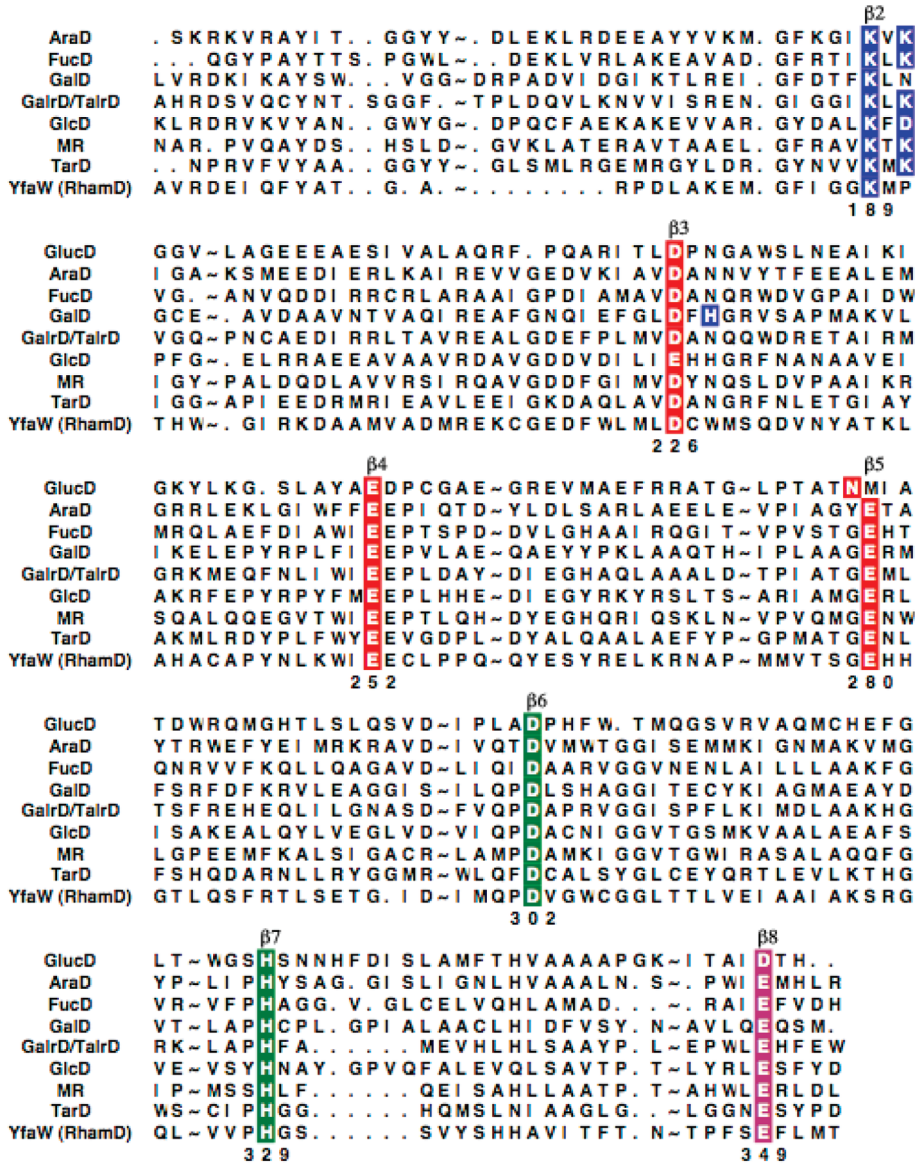
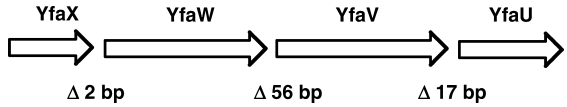


FIGURE 1: Alignment of the sequences of the $(\beta/\alpha)_7\beta$-barrel domains of members of the enolase superfamily that share the conserved His-Asp dyad located at the ends of the seventh and sixth β-strands. AraD, D-arabinonate dehydratase from *Sulfolobus solfarticus* (GI: 15899830); FucD, D-fuconate dehydratase from *Xanothomonas campestris* (GI: 21233491); GalD, D-galactonate dehydratase from *E. coli* K-12 (GI:49176390); GlcD, D-gluconate dehydratase from *Thermoproteus tenax* (GI:41033591); GlucD, D-glucarate dehydratase from *E. coli* K-12 (GI:16130694); MR, mandelate racemase from *Pseudomonas putida*; TalrD/GalrD, galactarate/L-talarate dehydratase from *S. typhimurium* LT2 (GI:16766982); TarD, D-tartrate dehydratase from *Bradyrhizobium japonicum* (GI:27381841); and YfaW (RhamD) from *E. coli* K-12 (GI:16130182). The identities of the conserved residues at the ends of the various β-strands are highlighted: the ligands for the Mg^{2+} are in red, the electrophilic and acid/base catalysts at the end of the second β-strand are in blue, the His-Asp dyad is in green, and the electrophilic carboxylate group is in magenta. In the sequence for GalD, the general acid/base catalyst at the end of the third β-strand is highlighted in blue. The numbers are those for the sequence of Yfaw (RhamD).

of the sequences of YfaW and other functionally assigned members of the MR and GlucD subgroups is displayed in Figure 1 (MR, AraD, GalD, GlcD, FucD, TalrD/GalrD, TarD, and GlucD). This alignment suggests that (1) the ligands for the Mg^{2+} in YfaW are Asp 226, Glu 252, and Glu 280, located at the C-terminal ends of the third, fourth, and fifth β-strands of the $(\beta/\alpha)_7\beta$-barrel domain; (2) His 329 and Asp 302 located at the C-terminal ends of the seventh and sixth β-strands, respectively, form the general basic His-Asp dyad; and (3) and Glu 349 located at the C-terminal end of the eighth β-strand is the electrophilic catalyst that hydrogen bonds with the carboxylate oxygen of the substrate that is not coordinated with the Mg^{2+}. The presence of these residues places YfaW in the MR subgroup and excludes it from the GlucD subgroup (Glu instead of Asn at the end of the fifth β-strand and Glu instead of Asp at the end of the eighth β-strand). In the MR subgroup, the α-OH acid substrates coordinate to the essential Mg^{2+} *via* the 2-OH group and one carboxylate oxygen.

In the structurally characterized MR (*13*), FucD (*7*), TarD (*11*), and TalrD/GalrD (*12*), two Lys residues are located at the end of the second β-strand: the first is an electrophilic catalyst that interacts with one carboxylate oxygen of the substrate and the corresponding enediolate oxygen of the intermediate, and the second is an acid/base catalyst. However, the sequence alignment also suggests that in YfaW only the electrophilic Lys residue is located at the end of the second β-strand, as was previously observed in the active site of GalD. In GalD, a His at the end of the third β-strand was identified as the acid catalyst that enables the anti-

Scheme 2



elimination reaction (*8*); a homologue of this His residue is absent in YfaW. Thus, the reaction catalyzed by YfaW contains a novel active site motif that enables a "new" function in the MR subgroup.

YfaW and its homologues[2] that share sufficient sequence identity to be considered isofunctional also share two His residues that are proximal to the active site (*vide infra*): one in the N-terminal capping domain (His 33 in the YfaW from *E. coli* K-12) and the second following Glu 280 at the end of the fifth β-strand (His 281). The presence of these residues raises the possibility that one or both might be an acid catalyst to facilitate departure of the 3-OH group. Thus, discovering both the identity of the reaction catalyzed by YfaW and its mechanism are important for understanding the structural bases for divergence of function in the MR subgroup.

Genome Context of YfaW and Its Orthologues. Based on the intergenic spacings, the gene encoding YfaW (GI: 16130182, annotated as a "putative racemase") appears to be located in an operon (*yfa*XWVU) that encodes three additional proteins [Scheme 2; the bp spacings between the open reading frames (Δ) are indicated]: YfaX (GI:16130183) annotated as a "putative (transcriptional) regulator", YfaV (GI:16130182) annotated as a "predicted transporter", and YfaU (GI:16130180) annotated as a "predicted 2,4-dihydroxyhept-2-ene-1,7-dioic acid aldolase".

YfaU is a member of a mechanistically diverse superfamily of Mg^{2+}-dependent enzymes that includes the structurally characterized 2-keto-3-deoxygalactarate aldolase (*26*), 4-hydroxy-2-oxo-heptane-1,7-dioate aldolase (HpcH) (*27*), and macrophomate synthase (*28*). The proximity of the genes encoding YfaW and YfaU, a possible 2-keto carboxylic acid aldolase, suggests that YfaW is an acid sugar dehydratase, not a racemase, and that its 2-keto-3-deoxy acid sugar product is the substrate for a retroaldol reaction catalyzed by YfaU (generating pyruvate and an aldehyde as products).

Not all of the presumed orthologues of YfaW share the same genome context (www.jgi.doe.gov; www.microbesonline.org). For example, orthologues of YfaW from strains of *Burkholderia cenocepacia* are encoded by operons that also encode proteins annotated as "L-rhamnose 1-epimerase" and "rhamosyltransferase" as well as an aldolase that is a member of the dihydropicolinate synthase superfamily. The latter enzymes utilize a Schiff base derived from pyruvate in their diverse reactions and mechanisms (*4, 29*). This genome context for a divergent RhamD supports the functional assignment of YfaW from library screening (next section).

Identification of L-Rhamnonate by Library Screening. YfaW from *E. coli* as well as its orthologue from *Salmonella typhimurium* LT2 (95% sequence identity) were screened for dehydration activity (formation of an α-keto acid product, semicarbazide reagent) using our library of mono- and diacid sugars. Four substrates were identified: L-rhamnonate (6-deoxy-L-mannonate), L-lyxonate, L-mannonate and D-gulonate (Scheme 3; these share the *lyxo*-configuration at carbons-2, -3 and -4.

Scheme 3

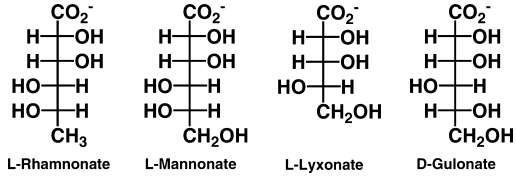


Table 2: Kinetic Constants for the RhamD Reaction[a]

organism	substrate	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})
E. coli K-12	L-rhamnonate	3.2 ± 0.2	0.15 ± 0.07	2.1×10^4
	L-lyxonate	0.6 ± 0.03	2.0 ± 0.3	3×10^2
	L-mannonate	0.2 ± 0.1	0.15 ± 0.06	1.3×10^3
	D-gulonate	0.02		
S. typhimurium LT2	L-rhamnonate	3.9 ± 1.6	0.25 ± 0.08	1.6×10^4
	L-lyxonate	0.05 ± 0.03	1.6 ± 1.2	3×10^1
	L-mannonate	0.1 ± 0.06	0.06 ± 0.04	1.7×10^3

[a] Assay conditions are given in Materials and Methods.

Table 3: Kinetic Constants for the 2-Keto-3-deoxy-L-rhamnonate Aldolase Reaction[a]

substrate	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})
2-keto-3-deoxy-L-rhamnonate	0.4 ± 0.1	0.078 ± 0.050	5.1×10^3
2-keto-3-deoxy-L-lyxonate	0.3 ± 0.01	0.8 ± 0.08	3.8×10^2
2-keto-3-deoxy-L-mannonate	0.3 ± 0.05	0.14 ± 0.03	2.1×10^3

[a] Assay conditions are given in Materials and Methods.

YfaW dehydrates L-rhamnonate with the greatest efficiency followed by L-lyxonate, L-mannonate and D-gulonate. The values of k_{cat} and k_{cat}/K_m for these substrates using YfaW from *E. coli* K-12 and its orthologue from *Salmonella typhimurium* LT2 (95% sequence identity, *vide infra*) are summarized in Table 2 and are comparable to those for other acid sugar dehydratases in the enolase superfamily.

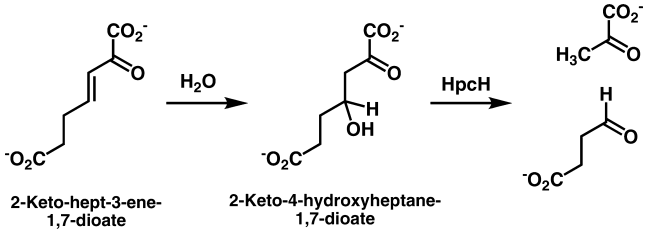
L-Rhamnonate, L-mannonate, and L-lyxonate share the same configurations at carbons-2, -3, and -4; D-gulonate differs from L-rhamnonate and L-mannonate by the configuration of carbon-5. As we observed in our studies of FucD (*7*), but not ManD (*5*) or TalrD/GalrD (*12*), YfaW is promiscuous with respect to the acid sugars that can be utilized as substrates.

L-Rhamnonate, L-lyxonate, and L-mannonate were dehydrated by YfaW, and the resulting 2-keto-3-deoxy acid sugars were used to determine whether YfaU catalyzes a retroaldol reaction as assessed by measuring the formation of pyruvate using NADH and LDH. At pH 8 and in the presence of Mg^{2+}, the dehydration products from all three acid sugars were converted to pyruvate, with 2-keto-3-deoxy-L-rhamnonate the best substrate based on the values of its kinetic constants (Table 3). Although the values of k_{cat} are less than those measured for homologous aldolases that utilize different substrates, the ability of YfaU to catalyze retroaldol reactions with the products of the YfaW-catalyzed reaction provides evidence that both participate in the same catabolic pathway.

Presumably, the substrate promiscuities that are observed in the dehydration and retroaldol reactions are "inadvertent" because these likely are unnatural compounds that would not be encountered by *S. typhimurium*. However, these

[2] http://sfld.rbvi.ucsf.edu.

Scheme 4



activities could allow selective advantage if these compounds were available for utilization as carbon sources.

The accompanying manuscript describes structural characterization of YfaU as well as the observation that it catalyzes a more efficient retroaldol reaction with 2-keto-4-hydroxyheptane-1,7-dioate (the substrate for the HpcH aldolase-catalyzed reaction; Scheme 4) than with 2-keto-3-deoxy-L-rhamnonate (*30*). Therefore, we examined whether YfaW also catalyzes the hydration of 2-keto-hept-3-ene-1,7-dioate to yield 2-keto-4-hydroxyheptane-1,7-dioate; no hydration was detected with either ecRhamD or stRhamD (data not shown).

As a result, we assign the L-rhamnonate dehydratase (RhamD) function to YfaW *E. coli* K-12 and its orthologues and the 2-keto-3-deoxy-L-rhamnonate aldolase function to YfaU and its orthologues.

Stereochemical Course of the RhamD-Catalyzed Reaction. The dehydration of L-rhamnonate catalyzed by the RhamD from *E. coli* K-12 (the former YfaW) was performed in both H_2O and D_2O so that the stereochemical course of the replacement of the 3-OH group by solvent-derived hydrogen could be determined. The dehydration product exists as an approximately equimolar mixture of α- and β-furanosyl hemiketals (Figure 2, panel A); the downfield proton associated with each 3-methylene group is preferentially labeled with deuterium (Figure 2, panels B and C).

The location of the solvent-derived deuterium (configuration of carbon-3) was determined by an NOE experiment (between the 6-methyl group and the diastereotopic protons of the 3-methylene group; data not shown). As noted in Figure 2, the proton spatially proximal to the methyl group is that replaced by deuterium, establishing that the dehydration reaction occurs with retention of configuration at carbon-3.

Structure of RhamD. Structures for unliganded RhamDs from *E. coli* K-12 (ecRhamD), *Salmonella typhimurium* LT2 (stRhamD), *Azotobacter vinlandii*, and *Gibberella zeae* were determined by the New York SGX Research Center for Structural Genomics (NYSGXRC). This manuscript reports structures of stRhamD in the presence of Mg^{2+} and, also, both Mg^{2+} and 3-deoxy-L-rhamnonate, an inert substrate analogue, that was obtained from the 2-keto-3-deoxy-L-rhamnonate dehydration product by reduction with $NaBH_4$ (Scheme 5). Representative electron density of the active site of stRhamD liganded with Mg^{2+} and 3-deoxy-L-rhamnonate is shown in Figure 3.

The polypeptide of RhamD possesses a bidomain structure, as expected for a member of the enolase superfamily (Figure 4, panel A). Residues 1−166 at the N-terminus and residues 392−405 at the C-terminus of the polypeptide form the α + β capping domain; residues 167−351 form the $(\beta/\alpha)_7\beta$-barrel domain, and residues 352−391 are located at the interface between the domains.

RhamD crystallizes as an octamer, in agreement with the oligomeric structure determined by gel filtration (data not shown). The octamer can be described as a tetramer of dimers, with the latter the asymmetric unit observed crystallographically. As in TalrD/GalrD (*12*), the minimum functional unit is a monomer because the active site is formed by residues from a single polypeptide. In MR (*13*), FucD (*7*), and TarD (*11*), the minimum functional unit is a dimer because one residue from the second polypeptide in the dimer completes the active site otherwise formed from the first polypeptide.

As observed in other members of the enolase superfamily, the binding site for the distal portion of the substrate is provided by the α + β capping domain. In MR and many of the structurally characterized members, the side chains of two loops form the substrate binding site (Figure 4, panel B): the first, designated the "20s loop", is ∼10 residues in length, connecting the first and second β-strands of a three strand antiparallel β-sheet at the N-terminus of the capping domain (highlighted in green); the second, designated the "50s loop", is also ∼10 residues in length and includes the C-terminal residues of the third β-strand of the same antiparallel β-sheet and the turn that connects to the first α-helix of the domain (highlighted in magenta). However, in RhamD, the "20s loop" is extended to 51 residues (18−68; also highlighted in green), and the third β-strand is reduced in length so that the "50s loop" does not participate in formation of the active site cavity (also highlighted in magenta). As a result, the binding site for the distal portion of the substrate is formed by the extended "20s loop".

By ^1H NMR spectroscopy, the reduced product was the expected mixture of two epimers at carbon-2 (Scheme 4); these could not be separated by ion-exchange chromatography, so the mixture was used for crystallization experiments. Polypeptide A in the dimeric asymmetric unit contained the epimer of this mixture in which carbon-2 has the same configuration as carbon-2 of L-rhamnonate, i.e., the inert 3-deoxy-L-rhamnonate epimer that mimics the substrate. As expected based on liganded structures of MR and TarD, the analogue is bound as a bidentate ligand of the Mg^{2+} involving the 2-OH and one oxygen of the carboxylate group; also, as expected, the second carboxylate oxygen is hydrogen-bonded to Glu 349 at the end of the eighth β-strand (Figure 5).

The second polypeptide (polypeptide B) contained a ligand that differed from that in polypeptide A; based on its electron density, it appears to be a 2,4-dihydroxyoctanoate anion, not one of the epimers of the reduced product. The origin of this molecule is unknown: it was not observed in crystals obtained in the absence of the epimeric mixture of reduced products; perhaps it can be explained by composite electronic density from 3-deoxy-L-rhamnonanate and L-methionine that also was present in the crystallization solution. Carbon-2 and -4 of this "ligand" share the same configurations as the analogous carbons in the substrate analogue present in polypeptide A, and the ligand is bound to the Mg^{2+} with the same bidentate geometry.

The Active Site of RhamD. The active site functional groups (Figure 5) are those predicted from the sequence alignment: (1) the ligands for the essential Mg^{2+}, Asp 226,

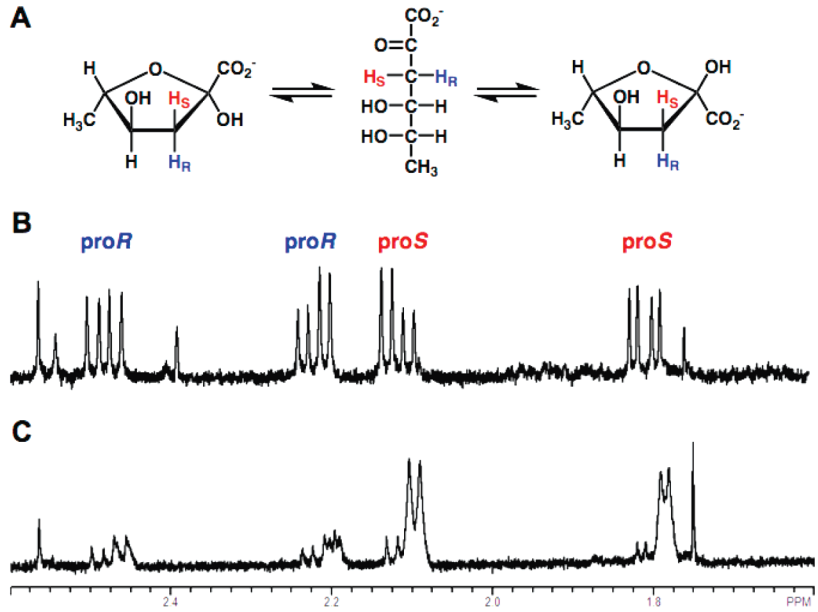
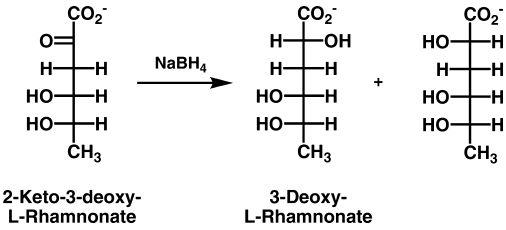


FIGURE 2: Panel A, the structures of the acyclic keto and cyclic hemiketal forms of the 2-keto-3-deoxy-L-rhamnonate product. Panel B, partial ^{1}H NMR spectrum of the product obtained by a reaction in H_2O showing the resonances associated with the 3-pro*R* and 3-pro*S* protons. Panel C, partial ^{1}H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

Scheme 5



Glu 252, and Glu 280 located at the C-terminal ends of the third, fourth, and fifth β-strands of the $(\beta/\alpha)_8$-barrel domain; (2) the His-Asp dyad is formed from His 329 and Asp 302 located at the ends of the seventh and sixth β-strands; and (3) in the structure of the analogue-liganded complex, Glu 349 is the electrophilic catalyst located at the end of the eighth β-strand. As suggested by the sequence alignment, an acid/base catalyst is not located at the end of the second β-strand; instead, Pro 191 is located at this position.

The active site contains the two additional His residues that are conserved in the RhamD family, His 33 in the extended 20s loop and His 281 sequence proximal to Glu 280, the Mg^{2+} ligand located at the end of the fifth β-strand. In the substrate analogue-liganded structure, O4 of the ligand is hydrogen bonded to Nε of His 33, suggesting that His 33 is conserved because this interaction is important for determining substrate specificity. His 281 is too distant from the analogue to be involved in either catalysis or substrate binding.

Therefore, based on the structure of the analogue-liganded complex, the dehydration reaction is almost certain to involve His 329 as the only acid/base catalyst, initially functioning as the base that abstracts the proton from carbon-2 to generate

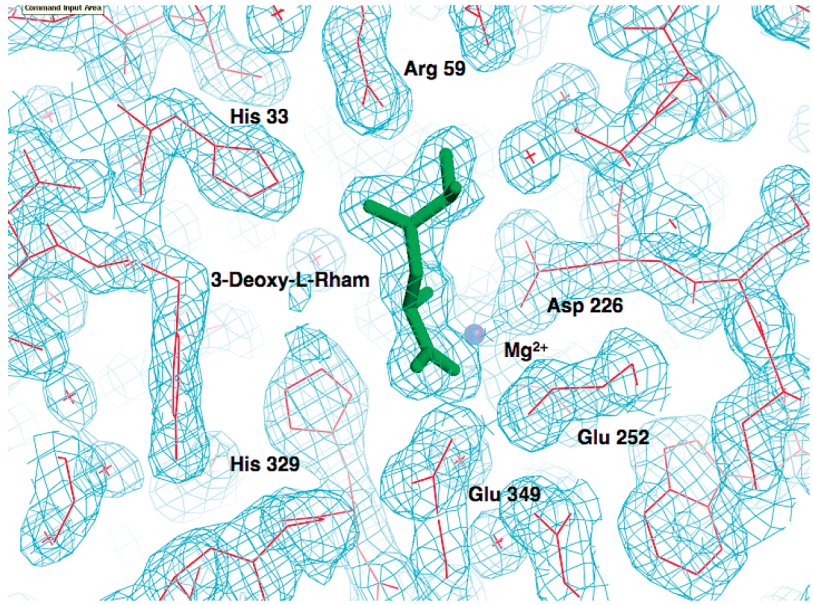


FIGURE 3: Representative electron density for the active site of stRhamD liganded with Mg^{2+} and 3-deoxy-L-rhamnonate contoured at 1.5σ. The details of the interactions between 3-deoxy-L-rhamnonate and the active site are described in the text.

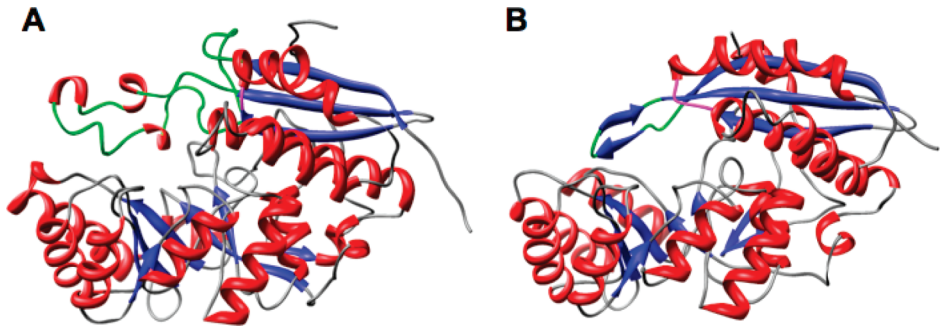


FIGURE 4: Structural comparison of stRhamD and MR. Panel A: In RhamD, the Gly 19 - Gly 68 loop connecting the C-terminal end of the first β-strand and the N-terminal end of the second β-strand in the capping domain is highlighted in green, and Ala 88 connecting the C-terminal end of the third β-strand and the N-terminal end of the first α-helix in the capping domain is highlighted in magenta. Panel B: In MR, the Pro 17 - Gly 30 loop ("20s loop") connecting the C-terminal end of the first β-strand and the N-terminal end of the second β-strand in the capping domain is highlighted in green, and the Ala 53−Thr 55 loop ("50s loop") connecting the C-terminal end of the third β-strand and the N-terminal end of the first α-helix in the capping domain is highlighted in magenta.

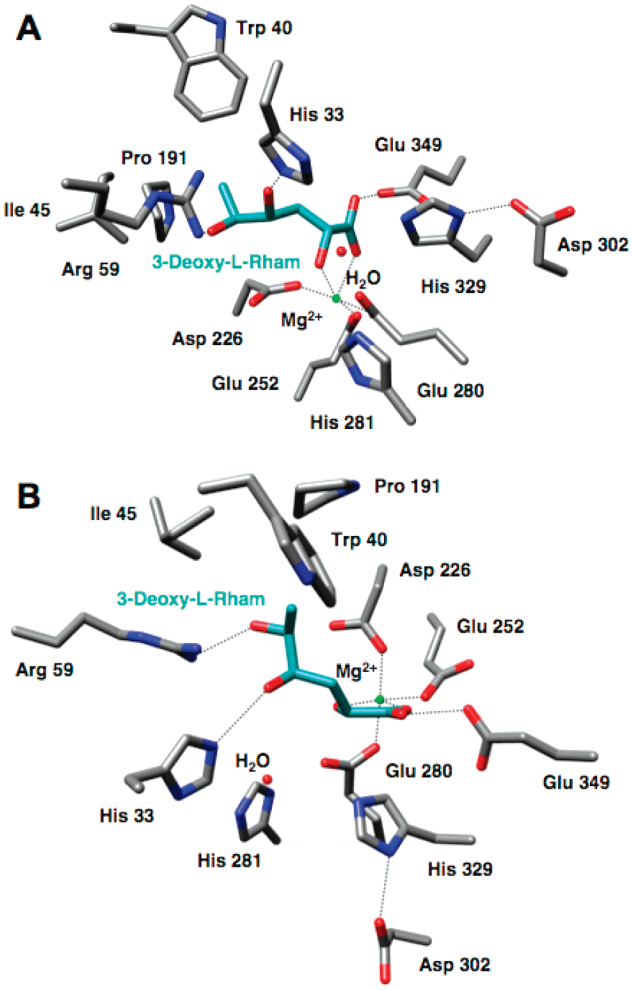


FIGURE 5: Views of the active site of stRhamD liganded with Mg^{2+} and 3-deoxy-L-rhamnonate. Panel A, "side" view highlighting the coordination of the essential Mg^{2+} and the disposition of the ligand to the His 329-Asp 302 dyad that functions as the only acid/base catalyst. Panel B, "top" view highighting the interactions of the distal portion of the ligand with active site hydrogen bond donors (O4 and O5) and hydrophobic residues (pocket for C6).

an enolate anion intermediate and subsequently as the acid that facilitates departure of the 3-OH group. When the enolate anion intermediate is formed, the change in hybridization of C2 from sp^3 to sp^2 will cause carbon-3 and its OH group to approach the protonated His 329, thereby allowing catalysis of the vinylogous elimination of the 3-OH group.

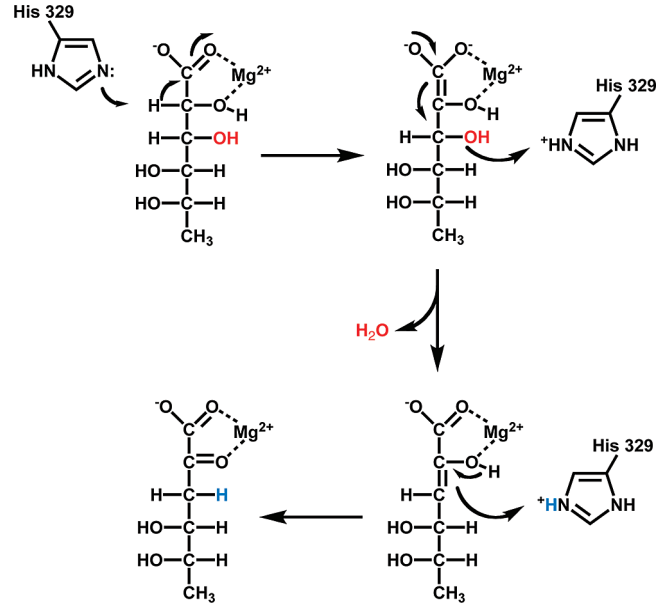


FIGURE 6: Proposed mechanism of the RhamD-catalyzed reaction. His 329, hydrogen-bonded to Asp 302, is positioned to function first as the base that abstracts the 2-proton to generate the stabilized enediolate intermediate. Both the subsequent *syn*-elimination of the 3-OH to generate an enol intermediate and its ketonization with overall retention of configuration to generate the product also are catalyzed by His 329.

Table 4: Kinetic Constants of Mutants for the RhamD Reaction[a]

enzyme	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})
wild type	3.2 ± 0.2	0.15 ± 0.07	2.1×10^4
H33N	—[b]	—	—
H281N	0.09 ± 0.01	8.8 ± 0.9	10
H329N	0.0005 ± 0.0001	0.3 ± 0.03	1.6

[a] Assay conditions are given in Materials and Methods. [b] No detectable activity.

Thus, RhamD catalyzes a *syn*-dehydration reaction. In addition, the replacement of the 3-OH group by solvent hydrogen with retention of configuration can be explained by participation of the conjugate acid of His 329 in this final step of the reaction (Figure 6).

Like the dehydration of D-glucarate by GlucD (*14*) and of L-talarate by TalrD/GalrD (*12*), the His of the His-Asp dyad is identified as the only acid/base catalyst in the RhamD-catalyzed reaction. However, in RhamD, the 2-OH group and one carboxylate oxygen of the ligand are ligands

of the essential Mg^{2+}; in GlucD, both carboxylate oxygens are ligands of the Mg^{2+}. So, despite the similarities in mechanism, the substrates are presented to the His catalysts with different geometries.

The structure of the analogue-liganded complex of RhamD contains a water molecule that is hydrogen-bonded to the carboxylate group of Glu 280 as well as the indole nitrogen of Trp 228. Although the water is not hydrogen-bonded to His 329, it may represent the position of the water that is eliminated.

In addition to His 33, the distal portion of the active site includes Arg 59 that also is part of the extended 20s loop (Figure 5, panel B). Its guanidinium group is also hydrogen-bonded to O5 of the substrate analogue, providing an additional specificity determinant. The remainder of the binding site for the distal portion of the substrate is hydrophobic (Trp 40, Ile 45, and Pro 191), consistent with the observed substrate preference for L-rhamnonate (6-deoxy-L-mannonate).

Kinetic Properties of Mutants of Active Site Residues. The structure of the analogue-liganded complex revealed the presence of three His residues in the active site, His 329, His 33, and His 281. Each of these was mutated to Asn ecRhamD (95% sequence identity) with significant negative impact on the kinetic constants (Table 4). As described in the previous section, only His 329 is appropriately positioned to participate in catalysis; His 33 appears to be important for substrate specificity, and His 281 may participate in the binding pocket for the water molecule that is eliminated. In the absence of structural information for the various mutants, quantitative interpretations of the observed activities are not possible.

Characterization of RhamD Orthologues. In parallel with our studies of ecRhamD and stRhamD, NYSGXRC purified orthologous RhamDs from *Azotobacter vinlandii, Gibberella zeae*, *Magnaporthe grisea*, *Polaromonas* sp. JS 666, and *Silicibacter pomeroyi.* Each of these homologues catalyzes the RhamD reaction, although the values of the k_{cat} and k_{cat}/K_m for some are less than those observed for the RhamDs from *E. coli* and *S. typhimurium* (data not shown).

NYSGXRC also determined unliganded structures of the RhamDs from *A. vinlandii* and *G. zeae*.[3] Not surprisingly, these share the same bidomain structure observed in ecRhamD and stRhamD as well as active site functional groups and specificity-determining residues.

Conclusions. The RhamD function has been assigned to YfaW from *E. coli* K-12 as well as isofunctional homologues that share active site residues we have implicated in both the mechanism of the dehydration reaction and determination of substrate specificity. The active site of RhamD contains a novel array of functional groups to catalyze a *syn*-dehydration reaction as well as a novel structure for the specificity determining loops in the capping domain. With the assignment of the RhamD function to YfaW, the only member of the enolase superfamily with unknown function encoded by the *E. coli* K-12 genome is the D-glucarate dehydratase-related-protein (GlucDRP) that shares 65% sequence identity with the GlucD that we have characterized both structurally and mechanistically (*19*).

[3] *A. vinlandii*, 2OZ3; *G. zeae*, 2P0I.

ACKNOWLEDGMENT

We thank Professor Christian P. Whitman, University of Texas at Austin, for materials and expert advice in assaying YfaW for 2-keto-hept-3-ene-1,7-dioate hydratase activity.

REFERENCES

1. 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*, 1159–1161.
2. 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.
3. 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.
4. 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*, 30591–30594.
5. Rakus, J. F., Fedorov, A. A., Fedorov, E. V., Glasner, M. E., Vick, J. E., Babbitt, P. C., Almo, S. C., and Gerlt, J. A. (2007) Evolution of enzymatic activities in the enolase superfamily: D-Mannonate dehydratase from Novosphingobium aromaticivorans. *Biochemistry 46*, 12896–12908.
6. Brouns, S. J., Walther, J., Snijders, A. P., van de Werken, H. J., Willemen, H. L., Worm, P., de Vos, M. G., Andersson, A., Lundgren, M., Mazon, H. F., van den Heuvel, R. H., Nilsson, P., Salmon, L., de Vos, W. M., Wright, P. C., Bernander, R., and van der Oost, J. (2006) Identification of the missing links in prokaryotic pentose oxidation pathways: evidence for enzyme recruitment. *J. Biol. Chem. 281*, 27378–27388.
7. 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*, 14582–14597.
8. 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.
9. 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*, 133–136.
10. 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*, 529–540.
11. 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*, 1459814–608.
12. 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*, 9564–9577.
13. 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-A resolution: identification of the active site and possible catalytic residues. *Biochemistry 30*, 9264–9273.
14. Gulick, A. M., Hubbard, B. K., Gerlt, J. A., and Rayment, I. (2001) Evolution of enzymatic activities in the enolase superfamily: identification of the general acid catalyst in the active site of D-glucarate dehydratase from Escherichia coli. *Biochemistry 40*, 10054–10062.
15. 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*, 486–491.
16. 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 Mg2+ and o-succinylbenzoate. *Biochemistry 39*, 10662–10676.

17. 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.
18. 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.
19. 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*, 4590–4602.
20. Burks, E. A., Jr., and Whitman, C. P. (1998) Sterochemical and isotopic labeling studies of 2-oxo-hept-4-ene-1,7-dioate hydratase: evidence for an enzyme-catalyzed ketonization step in the hydration reaction. *J. Am. Chem. Soc. 120*, 7665–7675.
21. 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.
22. Long, F., Vagin, A., Young, P., and Murshudov, G. N. (2008) BALBES: A molecular replacement pipeline. *Acta Crystallogr. D64*, 125–132.
23. Jones, A. T. (1985) Interactive computer graphics: FRODO. *Methods Enzymol. 115*, 157–171.
24. 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*, 905–921.
25. Lamzin, V. S., and Wilson, K. S. (1993) Automated refinement of protein models. *Acta Crystallogr. D49*, 129–147.
26. Izard, T., and Blackwell, N. C. (2000) Crystal structures of the metal-dependent 2-dehydro-3-deoxy-galactarate aldolase suggest a novel reaction mechanism. *EMBO J. 19*, 3849–3856.
27. Rea, D., Fulop, V., Bugg, T. D., and Roper, D. I. (2007) Structure and mechanism of HpcH: a metal ion dependent class II aldolase from the homoprotocatechuate degradation pathway of Escherichia coli. *J. Mol. Biol. 373*, 866–876.
28. Ose, T., Watanabe, K., Mie, T., Honma, M., Oikawa, H., and Tanaka, I. (2003) Insight into a natural Diels-Alder reaction from the structure of macrophomate synthase. *Nature 422*, 185–189.
29. Lawrence, M. C., Barbosa, J. A., Smith, B. J., Hall, N. E., Pilling, P. A., Ooi, H. C., and Marcuccio, S. M. (1997) Structure and mechanism of a sub-family of enzymes related to N-acetylneuraminate lyase. *J. Mol. Biol. 266*, 381–399.
30. Rea, D., Hovington, R., Rakus, J. F., Gerlt, J. A., Fülöp, V., Bugg, T. D. H., and Roper, D. I. (2008) Crystal structure and functional assignment of YfaU, a metal ion dependent class II aldolase from Escherichia coli K-12. *Biochemistry 47*, 9955–9965.

BI800914R