Evolution of Enzymatic Activities in the Enolase Superfamily: L-Fuconate Dehydratase from *Xanthomonas campestris*[†,‡]

Wen Shan Yew,[§] Alexander A. Fedorov,[||] Elena V. Fedorov,[||] John F. Rakus,[§] Richard W. Pierce,[§] 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, and Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461

Received August 18, 2006; Revised Manuscript Received September 18, 2006

ABSTRACT: Many members of the mechanistically diverse enolase superfamily have unknown functions. In this report we use both genome (operon) context and screening of a library of acid sugars to assign the L-fuconate dehydratase (FucD) function to a member of the mandelate racemase (MR) subgroup of the superfamily encoded by the *Xanthomonas campestris pv. campestris str.* ATCC 33913 genome (GI:21233491). Orthologues of FucD are found in both bacteria and eukaryotes, the latter including the rTS beta protein in *Homo sapiens* that has been implicated in regulating thymidylate synthase activity. As suggested by sequence alignments and confirmed by high-resolution structures in the presence of active site ligands, FucD and MR share the same active site motif of functional groups: three carboxylate ligands for the essential Mg^{2+} located at the ends of the third, fourth, and fifth β-strands in the $(\beta/\alpha)_7\beta$-barrel domain (Asp 248, Glu 274, and Glu 301, respectively), a Lys-x-Lys motif at the end of the second β-strand (Lys 218 and Lys 220), a His-Asp dyad at the end of the seventh and sixth β-strands (His 351 and Asp 324, respectively), and a Glu at the end of the eighth β-strand (Glu 382). The mechanism of the FucD reaction involves initial abstraction of the 2-proton by Lys 220, acid catalysis of the vinylogous β-elimination of the 3-OH group by His 351, and stereospecific ketonization of the resulting enol, likely by the conjugate acid of Lys 220, to yield the 2-keto-3-deoxy-L-fuconate product. Screening of the library of acid sugars revealed substrate and functional promiscuity: In addition to L-fuconate, FucD also catalyzes the dehydration of L-galactonate, D-arabinonate, D-altronate, L-talonate, and D-ribonate. The dehydrations of L-fuconate, L-galactonate, and D-arabinonate are initiated by abstraction of the 2-protons by Lys 220. The dehydrations of L-talonate and D-ribonate are initiated by abstraction of the 2-protons by His 351; however, protonation of the enediolate intermediates by the conjugate acid of Lys 220 yields L-galactonate and D-arabinonate in competition with dehydration. The functional promiscuity discovered for FucD highlights possible structural mechanisms for evolution of function in the enolase superfamily.

The discoveries that mandelate racemase (MR),[1] muconate lactonizing enzyme (MLE), and enolase share conserved amino acid sequences and three-dimensional structures first allowed the recognition that homologous enzymes need not catalyze the same chemical reaction (*1*); instead, homologous enzymes can catalyze different overall reactions that share a conserved partial reaction, in this case, Mg^{2+}-assisted enolization of a carboxylate anion substrate. Since these discoveries, the reactions catalyzed by several members of the enolase superfamily have been investigated in considerable mechanistic and structural detail (*2*). These include the ubiquitous enolase as well as other enzymes that participate in more specialized metabolic pathways. The latter enzymes include MR, D-glucarate dehydratase (GlucD), D-galactonate dehydratase (GalD), *o*-succinylbenzoate synthase, L-Ala-D/L-Glu epimerase (*3*), and, most recently, *N*-succinylamino acid racemase (*4*).

The members of the superfamily share a bidomain structure: an N-terminal "capping domain" that is primarily responsible for determining substrate specificity and a $(\beta/\alpha)_7\beta$-barrel domain that is primarily responsible for the acid/base chemistry (*5*). On the basis of structure−function studies that revealed the identities of positionally conserved ligands for the essential Mg^{2+} and the acid/base catalysts that both generate the enolate anion intermediate and direct it to product, the members of the enolase superfamily have been partitioned into three subgroups: enolase, MR, and MLE (*6*). The members of all three groups contain three conserved Asp/Glu residues individually located at the ends of the third, fourth, and fifth β-strands of the barrel domain. The members of the enolase subgroup always contain a general acid/base

[†] This research was supported by Grant GM-52594 and Program Project Grant GM-71790 from the National Institutes of Health.

[‡] The X-ray coordinates and structure factors for wild-type FucD liganded with Mg^{2+}, wild-type FucD liganded with Mg^{2+} and D-eryrthronohydroxamate, and the K220A mutant liganded with Mg^{2+} and L-fuconate have been deposited in the Protein Data Bank (PDB accession codes 2HNE, 2HXT, and 2HXU, respectively).

* Corresponding authors. 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.

[1] Abbreviations: FucD, L-fuconate dehydratase; GalD, D-galactonate dehydratase; GlucD, D-glucarate/L-idarate dehydratase; MLE, muconate lactonizing enzyme; MR, mandelate racemase; SD, saturation difference.

10.1021/bi061687o CCC: $33.50
Published on Web 11/14/2006

Lys located at the end of the sixth β-strand, the members of the MR subgroup always contain a general acid/base His-Asp dyad located at the ends of the seventh and sixth β-strands, respectively, and the members of the MLE subgroup always contain general acid/base Lys residues located at the ends of both the second and sixth β-strands.

Database searches reveal that all three subgroups are highly populated. Whereas the members of the enolase subgroup are considered to be isofunctional (i.e., all enolases) based on significant levels of sequence conservation, the members of the MR and MLE subgroups are considered to be functionally diverse. The latter conclusion is based both on their highly divergent sequences and on the observation that the majority of the members are bacterial and encoded by genes in different operon contexts. Thus, if the range of reactions catalyzed by the members of the superfamily is to be delineated so that the extent and limitations of divergent evolution can be understood, the functions of the divergent members of the MR and MLE subgroups must be determined. Unfortunately, the substrate specificities of the members of unknown function cannot yet be assigned on the basis of only sequence and structure, so their functions are unknown.

To both better understand the structural strategies for catalyzing divergent reactions as well as to facilitate the development of computational methods that might allow prediction of substrate specificity (*7*), we have adopted the strategy of screening unknown members with libraries of potential substrates to discover function. Recently, we reported the successful use of this approach to assign the *N*-succinylamino acid racemase function to members of the MLE subgroup (*4*). Although the sequences of members of the MR subgroup are diverse, many unknowns are encoded by genes in operons that also encode sugar kinases, dehydrogenases, aldolases, and/or mutarotases. Thus, the reasonable expectation is that these operons encode catabolic pathways for carbohydrates, with the unknown members of the enolase superfamily likely catalyzing the dehydration of acid sugars. In this report, we describe the first use of a library of acid sugars to discover the physiological function of an unknown member within the MR subgroup.

This report is focused on the assignment of function to a member of a large (>50 members) orthologous group of proteins (>50% sequence identity) in the MR subgroup that, based on sequence alignments, are predicted to contain a KxK motif at the end of the second β-strand in addition to the Asp, Glu, and Glu ligands for the essential Mg^{2+} at ends of the third, fourth, and fifth β-strands, respectively, a general acid/base His-Asp dyad located at the ends of the seventh and sixth β-strands, respectively, and a Glu at the end of the eighth β-strand. These residues are homologues of those found in the active site of MR; in the 1,1-proton transfer reaction catalyzed by MR, the second Lys in the KxK motif is the *S*-specific base, the His of the His-Asp dyad is the *R*-specific base, and the Glu residue at the end of the eighth β-strand participates in stabilization of the enediolate intermediate. In contrast to many unknown members of the enolase superfamily, sufficient operon context is present for some members of this orthologous group to provide independent evidence that supports the functional assignment made by library screening. Thus, these studies validate the screening approach for assigning function, thereby making it applicable to other unknown members for which operon context is insufficient and/or the products of other enzymes encoded by the operon are not readily available for functional characterization.

We used both operon context and library screening to independently assign the L-fuconate dehydratase (FucD) function to a member of this orthologous group from *Xanthomonas campestris* (GI:21233491); the orthologous group also includes the rTS beta protein in *Homo sapiens* that has been implicated in regulating thymidylate synthase activity. We conclude that (1) Lys 220 at the end of the second β-strand initiates the reaction by abstraction of the proton from the 2-carbon, (2) His 351 of the His 351-Asp 324 dyad is the general acid that facilitates departure of the 3-OH group, and (3) stereospecific delivery of a solvent-derived proton to generate the 2-keto-3-deoxy product occurs within the active site and may be catalyzed by Lys 220. Thus, the same constellation of functional groups in the $(\beta/\alpha)_7\beta$-barrel domain that catalyzes the 1,1-proton transfer reaction catalyzed by MR catalyzes the stereospecific anti-dehydration of L-fuconate in FucD, thereby highlighting the importance of substrate specificity determinants in determining the mechanism of the reaction. We also discovered that FucD can catalyze the slow epimerization and dehydration of substrate analogues that have the opposite configurations at carbon-2, revealing functional promiscuity that likely results from evolutionary processes that optimize the same active site architecture to catalyze different reactions.

MATERIALS AND METHODS

^{1}H NMR spectra were recorded using either a Varian Unity INOVA 750 MHz or a 500NB MHz NMR spectrometer.

Library of Acid Sugars. A description of the synthesis of mono- and diacid sugars which were not commercially available can be found in the Supporting Information. The monoacids which were synthesized by Br_2 oxidation of the corresponding aldose and purified by anion-exchange chromatography using a Dowex AG1-X8 column eluted with a gradient of formic acid; the lactones that were obtained following rotary evaporation of the solvent were hydrolyzed by raising the pH to 10 with NaOH and, when the reaction was complete, neutralization to pH 7 with HCl. The diacids were synthesized either by nitric acid oxidation of the aldose or alditol or by 4-acetamido-TEMPO oxidation of the aldose and purified by anion-exchange chromatography using a Dowex AG1-X8 column; the lactones that were obtained following rotary evaporation of the solvent were also hydrolyzed by raising the pH to 10 with NaOH and, when the reaction was complete, neutralization to pH 7 with HCl.

Cloning, Expression, and Protein Purification of L-*Fuconate Dehydratase, a Protein Previously Annotated at the RTS Beta Protein*. The gene encoding RTS beta protein (GI:21233491) was PCR amplified from genomic DNA isolated from *X. campestris pv. campestris str.* ATCC 33913 (ATCC) using platinum *Pfx* DNA polymerase (Invitrogen). The PCR reaction (100 μL) contained 1 ng of plasmid DNA, 10 μL of 10× *Pfx* amplification buffer, 1 mM $MgSO_4$, 0.4 mM dNTP, 40 pmol of each primer (forward primer 5′-CCAGGACTCATATGCGCACCATCATCGC-CCTCGAGACC-3′ and reverse primer 5′-GCAGCACGCG-GTGCGTTGAGGATCCTTAGGCCTTCGCC-3′), and 5 units

of platinum *Pfx* DNA polymerase. The gene was amplified using 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 temperature range of 45−60 °C for 1 min and 15 s, 68 °C for 2 min, and a final extension of 68 °C for 10 min. The amplified gene was cloned into a modified pET-15b (Novagen) vector in which the N-terminal His tag contains 10 instead of the usual 6 His residues.

The protein was expressed in *Escherichia coli* strain BL21-(DE3). Transformed cells were grown at 37 °C in LB broth (supplemented with 100 μg/mL ampicillin) for 48 h and harvested by centrifugation. No IPTG was added to induce protein expression. The cells were resuspended in binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9, and 5 mM $MgCl_2$) and lysed by sonication. The lysate was cleared by centrifugation, and the His-tagged protein was purified using a column of chelating Sepharose Fast Flow (Amersham Biosciences) charged with Ni^{2+}. Cell lysate was applied to the column in binding buffer, washed with 15% elute buffer (1 M imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9, and 5 mM $MgCl_2$)−85% wash buffer (60 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9, and 5 mM $MgCl_2$), and eluted with 50% binding buffer−50% strip buffer (100 mM EDTA, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9, and 5 mM $MgCl_2$). The N-terminal His tag was removed with thrombin (Amersham Biosciences) according to the manufacturer's instructions, and the proteins were purified to homogeneity on a Q-Sepharose high-performance column (Amersham Biosciences) equilibrated with binding buffer (25 mM Tris-HCl, pH 7.9, and 5 mM $MgCl_2$) and eluted with a linear gradient of 0−0.5 M elution buffer (1 M NaCl, 25 mM Tris-HCl, pH 7.9, and 5 mM $MgCl_2$).

*Cloning, Expression, and Protein Purification of 2-Keto-3-deoxy-*L*-fuconate 4-Dehydrogenase, a Protein of Previously Unknown Function*. The gene encoding oxidoreductase (GI:21233489) was PCR amplified from genomic DNA isolated from *X. campestris pv. campestris str.* ATCC 33913 (ATCC) using platinum *Pfx* DNA polymerase (Invitrogen), the forward primer 5′-CAGCGCGCGGACTACGCCATATG-GCTGCAGATCGCACAGC-3′, and the reverse primer 5′-C T T G A T G C G T T G G G A T C C T C A G T T G G A C-CAGCCGCCGTCG-3′. Conditions for cloning, expression ,and protein purification were identical to those of the RTS beta protein previously described.

Cloning, Expression, and Protein Purification of L*-Fucopyranoside Mutarotase, a Protein of Previously Unknown Function*. The gene encoding L-fucopyranoside mutarotase (GI:21233492) was PCR amplified from genomic DNA isolated from *X. campestris pv. campestris str.* ATCC 33913 (ATCC) using platinum *Pfx* DNA polymerase, the forward primer 5′-GGTGGCGGGCATATGAGCATG-CAGCGGCTGTGCTATGTGC-3′, and the reverse primer 5′-CTACACAGTACGACTCGAGCTACGCAGCA-GAACCCTGCTGC-3′. Conditions for cloning, expression, and protein purification were identical to those of the RTS beta protein previously described.

Synthesis of D*-Erythronohydroxamate.* An excess volume of 50% aqueous hydroxylamine was added to a solution of D-erythrono-1,4-lactone (1 mmol), and the solution was allowed to stir for 10 min at room temperature to yield D-erythronohydroxamic acid. Unreacted hydroxylamine and solvent were removed by rotary evaporation. The product was characterized by ^1H and ^{13}C NMR spectroscopy. ^1H NMR (500 MHz, D_2O): δ 4.04 (d, J = 5.3 Hz, 1H), 3.78 (m, 1H), 3.56 (dd, J = 3.7 and 11.9 Hz, 1H), 3.5 (dd, J = 6.9 and 11.9 Hz, 1H). ^{13}C NMR (500 MHz, D_2O): δ 171.0 (C1), 73.0 (C2), 71.8 (C3), 62.2 (C4).

Site-Directed Mutagenesis and Protein Purification. The site-directed K220A mutant of FucD was constructed using the QuikChange kit (Stratagene), verified by sequencing, expressed in the BL21(DE3) *E. coli* cells, and purified as previously described for wild-type FucD.

Screening a Library of Acid Sugars for Dehydratase Activity. Dehydration of the members of a library of mono- and diacid sugars was monitored after 16 h by end-point detection of the semicarbazone derivative at 250 nm. The dehydration reactions were performed at 30 °C in 50 mM Tris-HCl (pH 8.0), containing 10 mM $MgCl_2$, 10 mM acid sugar, and 1 μM enzyme. A list of the acid sugars used, including the details of synthesis for those not available commercially, is included in the Supporting Information.

L-Fuconate, the proposed physiological substrate of FucD, was synthesized by Br_2 oxidation of L-fucose (Sigma) followed by purification by anion-exchange chromatography using a Dowex AG1-X8 column eluted with a gradient of formic acid; the lactone that was obtained following rotary evaporation of the solvent was hydrolyzed by raising the pH to 10 with NaOH and, when the reaction was complete, neutralization to pH 7.5 with HCl. ^1H NMR (500 MHz, D_2O): δ 4.08 (d, J = 1.6 Hz, 1H, C2), 3.92 (qd, J = 1.9 and 6.6 Hz, 1H, C5), 3.75 (dd, J = 1.6 and 9.5 Hz, 1H, C3), 3.27 (dd, J = 1.9 and 9.5 Hz, 1H, C4), 1.07 (d, J = 6.6 Hz, 3H, C6).

Screening a Library of Acid Sugars for Epimerization and Exchange of the α*-Proton.* To determine whether FucD catalyzes either exchange of the α-proton and/or epimerization of the members of the library of acid sugars, ^1H NMR spectra were recorded after FucD was incubated for 16 h with the members of the library in a D_2O-containing buffer at 25 °C. A typical reaction (800 μL) contained a mixture of pairs of acid sugars (to reduce the number of NMR spectra) at a concentration of 2 mM for each acid sugar, 50 mM Tris-DCl, pD 7.5, 10 mM $MgCl_2$, and 1 μM FucD. The rate of exchange (k_{exc}) of the α-proton of various monocarboxylate acid sugars was determined as described for L-fuconate. FucD was exchanged into D_2O buffer (50 mM Tris-DCl, pD 7.5) using an Amicon (10000 Da) stirred ultrafiltration cell by repeated concentration and dilution.

Crystallization and Data Collection. All crystals, wild-type FucD liganded with Mg^{2+}, selenomethionine-substituted FucD liganded with Mg^{2+}, wild-type FucD liganded with Mg^{2+} and D-erythronohydroxamate, and the K220A mutant liganded with Mg^{2+} and L-fuconate, were grown by the hanging drop method at room temperature. The crystallization conditions for the first two samples utilized precipitant composed of 20% PEG 3350 and 100 mM magnesium formate, pH 5.9. For the third sample, wild-type FucD liganded with Mg^{2+} and D-erythronohydroxamate, the precipitant was 56% tacsimate, pH 7.0. For the fourth case, the K220A mutant liganded with Mg^{2+} and L-fuconate, the precipitant was 2 M $(NH_4)_2SO_4$ and 100 mM Hepes, pH 7.5. The protein solution for the first sample contained wild-type FucD (30 mg/mL) in 50 mM Tris-HCl, pH 7.9, containing

Table 1: Data Collection and Refinement Statistics

				SeMet FucD·Mg		
	WT FucD·Mg	WT FucD·Mg·D-EHM	K220A FucD·Mg·L-Fuc	edge	peak	remote
data collection						
beamline	APS SGX-CAT	NSLS X29	NSLS X4A	NSLS X9A		
wavelength	0.979	0.979	0.979	0.97934	0.97911	0.97166
space group	$P3_221$	$P622$	$P622$	$P3_221$		
unit cell parameters						
a (Å)	130.63	159.53	159.15	130.52		
c (Å)	195.82	102.06	101.96	193.74		
resolution (Å)	2.0	1.7	1.8	2.34	2.34	2.34
unique reflections	122876	82705	67796	78296	78328	78301
completeness (%)	94.1	98.5	96.2	96.5	96.8	96.8
R_{merge}	0.074	0.097	0.091	0.095	0.075	0.109
refinement						
resolution (Å)	25.0−2.0	25.0−1.7	25.0−1.8	25.0−2.34		
R_{cryst}	0.241	0.179	0.172	0.247		
R_{free}	0.274	0.197	0.183	0.272		
rmsd, bonds (Å)	0.006	0.005	0.005	0.007		
rmsd, angles (deg)	1.40	1.34	1.31	1.48		
no. of protein atoms	13320	3347	3343	13315		
no. of ligand atoms		10	12			
no. of Mg^{2+}	4	1	1	4		
no. of waters	295	426	355	205		
PDB ID	2HNE	2HXT	2HXU	1YEY		

100 mM NaCl and 5 mM $MgCl_2$. The protein solution for the second sample contained selenomethionine-substituted FucD (42 mg/mL) in 50 mM Tris-HCl, pH 7.9, containing 100 mM NaCl, 5 mM $MgCl_2$, and 1 mM β-mercaptoethanol. The protein solution for the third sample contained wild-type FucD (42 mg/mL) in 50 mM Tris-HCl, pH 7.9, containing 100 mM NaCl, 5 mM $MgCl_2$, and 60 mM D-erythronohydroxamate. And, the protein solution for the fourth sample contained the K220A mutant (42 mg/mL) in 50 mM Tris-HCl, pH 7.9, containing 100 mM NaCl, 5 mM $MgCl_2$, and 60 mM L-fuconate. In the first two samples crystals appeared in 2 days and exhibited diffraction consistent with space group $P3_221$, with four molecules of the binary complex per asymmetric unit. For the last two samples crystals appeared in 5 days and exhibited diffraction consistent with the space group $P622$ with one molecule of the ternary complex per asymmetric unit. Prior to data collection the crystals of the wild-type FucD liganded with Mg^{2+} and the selenomethionine-substituted FucD liganded with Mg^{2+} were transferred to a cryoprotectant solution composed of 20% PEG 3350, 100 mM magnesium formate, pH 5.9, and 20% glycerol and flash-cooled in a nitrogen stream. The crystals of the wild-type FucD liganded with Mg^{2+} and D-erythronohydroxamate were transferred to a cryoprotectant solution composed of 56% tacsimate, pH 7.0, 5 mM $MgCl_2$, 60 mM D-erythronohydroxamate, and 30% glycerol. The crystals of the K220A mutant liganded with Mg^{2+} and L-fuconate were transferred to a cryoprotectant solution composed of 2 M $(NH_4)_2SO_4$, 0.1 M Hepes, pH 7.5, 5 mM $MgCl_2$, 60 mM L-fuconate, and 30% glycerol. Three-wavelength MAD data sets for SeMet FucD were collected to 2.34 Å resolution at the NSLS X9A beamline (Brookhaven National Laboratory) on a MarCCD-165 detector. X-ray diffraction data sets for wild-type FucD liganded with Mg^{2+} were collected to 2.0 Å at the SGX-CAT beamline at the Advanced Photon Source; for wild-type FucD liganded with Mg^{2+} and D-erythronohydroxamate data were collected to 1.7 Å at the NSLS X29A beamline; for the K220A FucD mutant liganded with Mg^{2+} and L-fuconate data were collected to 1.8 Å at the NSLS X4A beamline on an ADSC CCD detector. Diffraction intensities were integrated and scaled with programs DENZO and SCALEPACK (*8*). The data collection statistics are given in Table 1.

Structure Determination and Refinement. Initial attempts to determine the structure of the wild-type FucD with molecular replacement using other members of enolase superfamily as search models were unsuccessful. The structure of SeMet FucD was solved by multiple anomalous dispersion (MAD) with the program SOLVE (*9*) where 33 out of 36 selenium sites were identified. These heavy atom sites were used to calculate initial phases which were improved by solvent flattening and CNS averaging with the program RESOLVE (*10*), yielding an interpretable map for all four molecules in the asymmetric unit. Iterative cycles of manual rebuilding with TOM (*11*) and refinement with CNS resulted in a model with $R_{cryst} = 0.247$ and $R_{free} = 0.272$ at 2.34 Å resolution. This model was refined subsequently at 2.0 Å for wild-type FucD to $R_{cryst} = 0.241$ and $R_{free} = 0.274$. Four molecules in the asymmetric unit of the wild-type FucD structure form two tight dimers, typical for members of the enolase superfamily. The Mg^{2+} ions were clearly visible in the electron density maps of wild-type FucD and were coordinated by Asp 248, Glu 274, Glu 301, and three water molecules. A short chain segment in the 20's loop in two of the four monomers in the asymmetric unit exhibited weak density, and these residues were not included in the final model of wild-type FucD. The structures of two ternary complexes, wild-type FucD·Mg^{2+}·D-erythronohydroxamate and K220A FucD·Mg^{2+}·L-fuconate, were solved by molecular replacement with the program EPMR, using the wild-type FucD monomer as the search model. The bound ligands and Mg^{2+} were clearly visible already in the initial electron density maps in both structures. Iterative cycles of automatic rebuilding with ARP, manual rebuilding with TOM, and refinement with CNS were carried out in both cases. All residues of the protein chain are well-defined and are included in the final models of both structures. The structure of the ternary wild-type FucD·Mg^{2+}·D-erythrono-

hydroxamate complex was refined at 1.7 Å with R_{cryst} = 0.179 and R_{free} = 0.197. The structure of the ternary K220A FucD·Mg^{2+}·L-fuconate complex was refined at 1.8 Å with R_{cryst} = 0.172 and R_{free} = 0.183. Final refinement statistics are given in Table 1.

Dehydration of L*-Fuconate Monitored by* ^{1}*H NMR Spectroscopy.* Dehydration of L-fuconate by FucD was determined by ^{1}H NMR spectroscopy. The sample for NMR analysis (800 μL at 25 °C) contained 10 mM L-fuconate, 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, and 1 μM FucD. The mixture was incubated for 16 h at 25 °C, lyophilized, and resuspended in 800 μL of D_2O. A duplicate sample without enzyme was prepared for comparison.

Kinetic Assay of L*-Fuconate Dehydratase (FucD).* FucD activity was assayed by a continuous coupled spectrophotometric assay, using 2-keto-3-deoxy-L-fuconate dehydrogenase (previously annotated as oxidoreductase, cloned and expressed from the same operon containing FucD) to reduce the dehydrated 2-keto-3-deoxy-L-fuconate product. The assay (1 mL at 25 °C) contained L-fuconate, 50 mM K^+Hepes, pH 7.5, 10 mM $MgCl_2$, 1.5 mM NAD^+, 1.5 mM *p*-iodonitrotetrazolium violet (INT), 5 units of diaphorase, and 10 units of 2-keto-3-deoxy-L-fuconate dehydrogenase. The conversion of INT from the oxidized to the reduced form at 25 °C was monitored by the increase in absorbance at 500 nm; the molar extinction coefficient of reduced INT was determined to be 12990 M^{-1} cm^{-1} using diaphorase to oxidize known concentrations of NADH.

*Kinetic Assay of 2-Keto-3-deoxy-*L*-fuconate 4-Dehydrogenase.* 2-Keto-3-deoxy-L-fuconate 4-dehydrogenase activity was assayed by a continuous, coupled-enzyme spectrophotometric assay. The assay (1 mL at 25 °C) contained 2-keto-3-deoxy-L-fuconate (prepared in situ with L-fuconate and FucD), 50 mM K^+Hepes, pH 7.5, 10 mM $MgCl_2$, 1.5 mM NAD^+, 1.5 mM *p*-iodonitrotetrazolium violet (INT), and 5 units of diaphorase. The conversion of INT from the oxidized to the reduced form at 25 °C was similarly monitored by the increase in absorbance at 500 nm, with the molar extinction coefficient of reduced INT determined to be 12990 M^{-1} cm^{-1}.

Kinetic Assay of L*-Fucopyranoside Mutarotase Activity.* L-Fucopyranoside mutarotase activity was assayed using the saturation difference (SD) ^{1}H NMR technique (*12*, *13*). The assay (800 μL at 20 °C) contained L-fucose (5−100 mM), 50 mM sodium phosphate, pD 7.5, 100 mM NaCl, and 25 μM L-fucopyranoside mutarotase. The first and second free induction decays (FIDs) were obtained by presaturating the selected peak (the H1 peak of the α-anomer) in the first FID and shifting the receiver phase 180° for the second FID. The intensity of the saturated peak was used as a concentration reference standard. For analyzing enzymatic constants, the forward conversion rate ($k^A{}_2$[EA]) mediated by the mutarotase can be simplified by the Haldane relation at equilibrium ($k^A{}_2$[EA] = $k^B{}_2$[EB]) to the equation:

$$k^A{}_2[\mathrm{EA}] = (k_{T_1}[\mathrm{B}]F_P)/(1 - F_P)$$

where k_{T_1} is the relaxation decay constant determined from T_1-relaxation experiments, [B] is the equilibrium concentration of the β-anomer, and F_P is the ratio of the concentration of the saturation transferred peak (determined from the SD experiments) over the equilibrium concentration of the unsaturated pyranoside peak [B].

FucD-Catalyzed Exchange of the α*-Proton of* L*-Fuconate.* L-Fuconate was incubated with FucD in D_2O buffer at 20 °C, and ^{1}H NMR spectra were recorded as a function of time. A typical reaction contained 10 mM L-fuconate, 50 mM Tris-DCl (pD 7.5), 10 mM $MgCl_2$, and 0.5 μM FucD. The rate of exchange (k_{exc}) of the α-proton of L-fuconate was calculated from eq 1, which takes into account the total amount of L-fuconate bound at any time.

$$k_{exc} = k_{obs}[\text{L-fuconate}]_T/[\text{L-fuconate}]_B \qquad (1)$$

where k_{obs} is the observed first-order rate constant, [L-fuconate]$_T$ is the total L-fuconate concentration, and [L-fuconate]$_B$ is the concentration of bound L-fuconate. Because the concentration of bound L-fuconate (10 mM) was significantly higher than the enzyme concentration (0.5 μM), the concentration of bound L-fuconate ([L-fuconate]$_B$) was equal to the enzyme concentration.

Stereochemical Course of L*-Fuconate Dehydration.* To determine the stereochemical course of dehydration, L-fuconate was dehydrated in a D_2O-containing buffer at 20 °C, with ^{1}H NMR spectra recorded upon completion of the reaction (as assayed by semicarbazone formation). A typical reaction (800 μL) contained 10 mM L-fuconate, 50 mM Tris-DCl, pD 7.5, 10 mM $MgCl_2$, and 1 μM FucD. FucD was exchanged into D_2O buffer (50 mM Tris-DCl, pD 7.5) using an Amicon (10000 Da) stirred ultrafiltration cell.

RESULTS AND DISCUSSION

Library of Acid Sugars for Screening. As noted in the introduction, many members of the MR subgroup are encoded by operons that also encode sugar kinases, dehydrogenases, aldolases, and/or mutarotases. On the basis of this genomic context, our expectation is that the members of the MR subgroup encoded by these operons are acid sugar dehydratases or acid sugar epimerases; we also expect that other members of the MR subgroup that are not encoded by operons also utilize acid sugars as substrates. Some acid sugars are commercially available, e.g., D-glucarate, galactarate, D-galacturonate, and D-glucuronate; others are available as the aldonolactones, e.g., D-galactonate and D-mannonate.

Although the metabolites tabulated in the KEGG database (Kyoto Encyclopedia of Genes and Genomes; http://www.genome.ad.jp/kegg/ligand.html) include a variety of mono- and diacid sugars, we expect that new metabolic pathways remain to be discovered, and unknown members of MR subgroup will participate in these. Therefore, we decided to assemble a library of potential substrates that includes *all* aldonic and aldaric acids derived from hexoses, pentoses, and tetroses (Figure 1). We recognize that this library will be insufficient to assign function to all of the unknown members: indeed, in unpublished experiments we have failed to identify the functions of several unknown members using this library. We continue to add new acid sugars to our library as clues, time, and resources permit.

As described in this report, our experimental approach is to first screen members of the MR subgroup for dehydration activity using a semicarbazide end-point measurement; we use large concentrations of the unknown enzyme (typically

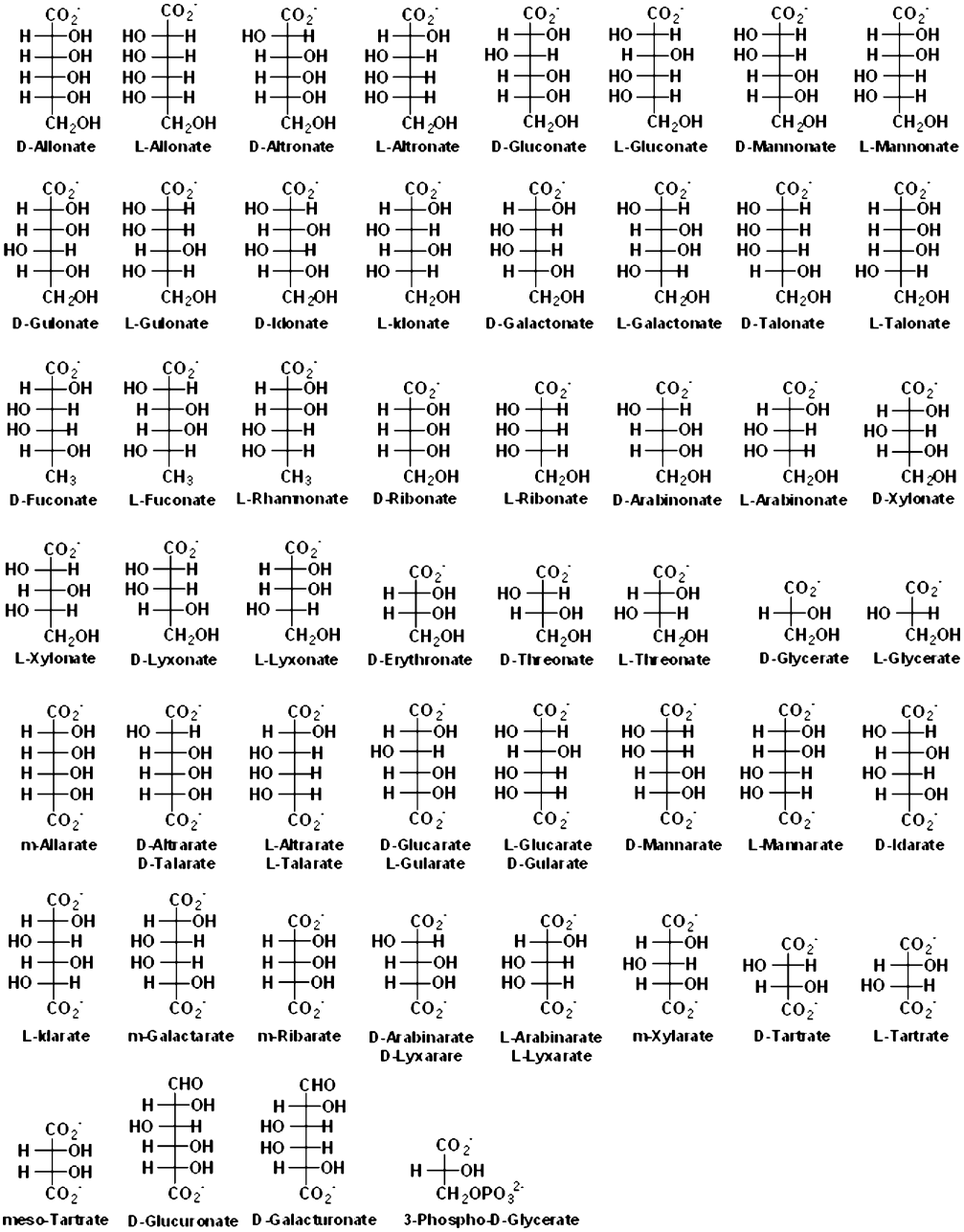


FIGURE 1: Library of acid sugars.

1 μM) and long incubation times (16 h) to allow partial reactions to be discovered, because these may provide clues as to the identity of the physiological substrate. Because we expect that some unknown members could function as epimerases and/or dehydratases in catabolic pathways, we also use our library to screen for epimerization reactions by incubating the members of the library with the unknown enzyme in D_2O-containing buffers and monitoring the reactions with ^1H NMR spectroscopy. As described in this report, the screen for epimerization also allows us to assess whether the unknown enzyme can catalyze exchange of the α-proton of the substrate, thereby providing additional clues as to the identity of the physiological substrate.

Operon Context and Proposed Function of XCC 4069 from X. campestris pv. campestris str. ATCC 33913 (GI: 21233491). A member of the MR subgroup encoded by the genome of *X. campestris* (GI:21233491; XCC 4069) predicted to have a Lys general acid/base catalyst at the end of the second β-strand and annotated as "RTS beta protein" was selected for functional assignment based on compelling information about substrate specificity obtained from analyses of the likely functions of other proteins encoded by the same operon. In *H. sapiens*, the RTS beta protein has been implicated in regulating thymidylate synthase activity (*14*−*16*).

The putative operon is displayed in Figure 2A. On the basis of the expected functions of homologues of the proteins encoded by each of the genes in this operon (as identified with BLASTP and PSI-BLAST), we proposed a novel catabolic pathway for the catabolism of L-fucose by *X. campestris* (Figure 2B):

(1) XCC 4071 (GI:21233493) is annotated as "fucose permease" in the protein sequence database. Several orthologues are annotated as "L-fucose permease," based on the assignments of function to the genes in the L-fucose catabolic pathway/operon in *E. coli* K-12 (*17*, *18*). Thus, on the basis

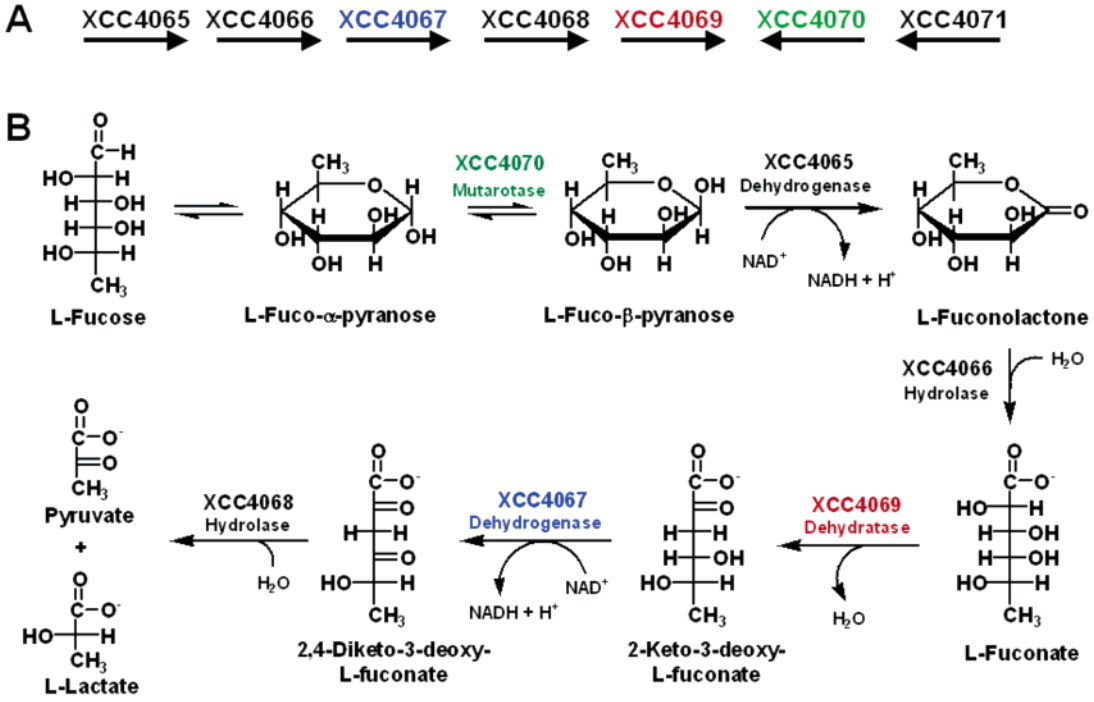


FIGURE 2: (A) Operon in the *X. campestris* genome that encodes the MR homologue XCC 4069. (B) Pathway for the utilization of L-fucose. The genes highlighted in color correspond to proteins that have been successfully purified and functionally assigned.

of this assignment, we postulated that the operon encoding the RTS beta protein would encode a catabolic pathway for L-fucose, although the L-fucose catabolic pathway in *E. coli* does not include a member of the enolase superfamily (L-fucose is converted to L-fuculose 1-phosphate, which is cleaved to yield dihydroxyacetone phosphate and lactaldehyde for further oxidative degradation; http://www.ecocyc.org/).

(2) XCC 4070 (GI:21233492) is annotated as a "hypothetical protein". PSI-BLAST searches of the databases disclose a weak homology with L-rhamnose mutarotase that equilibrates the α- and β-anomers of L-rhamnose (*19*). Recognizing the L-fucose also exists in solution as a mixture of slow-converting α- and β-anomers, XCC 4070 was proposed to be L-fucopyranoside mutarotase.

(3) XCC 4065 (GI:21233487) is annotated as an "oxidoreductase". BLASTP searches of the databases reveal that XCC 4065 is a homologue of aldose ketoreductases, including some that oxidize aldoses to aldonolactones. We postulated that XCC 4065 catalyzes the oxidation of an anomer of L-fucopyranoside to L-fuconolactone.

(4) XCC 4066 (GI:21233488) is annotated as a hypothetical protein. BLASTP searches reveal that XCC 4066 is a homologue of metal-dependent amidohydrolases. We postulated that XCC 4066 is L-fuconolactone hydrolase that catalyzes the conversion of L-fuconolactone to L-fuconate.

(5) XCC 4069 (GI:21233491), the MR homologue, is annotated as RTS beta protein. The RTS beta protein is a human protein of unknown function that is associated with control of the expression of thymidylate synthase (*14*−*16*). Overexpression of this protein downregulates not only intracellular levels of thymidylate synthase but also cells that are not in direct cellular contact, suggesting the involvement of a diffusable signaling molecule. We postulated that XCC 4069 is L-fuconate dehydratase that catalyzes the conversion of L-fuconate to 2-keto-3-deoxy-L-fuconate.

(6) XCC 4067 (GI:21233489) is annotated as an oxidoreductase. BLASTP searches of the databases reveal that XCC 4066 is a member of the short chain dehydrogenase/reductase superfamily; many members of this superfamily catalyze the oxidation of secondary alcohols. We postulated that XCC 4069 is 2-keto-3-deoxy-L-fuconate 4-dehydrogenase that catalyzes the NAD^+-dependent oxidation of 2-keto-3-deoxy-L-fuconate to 2,4-diketo-3-deoxy-L-fuconate.

(7) XCC 4068 (GI:21233490) is annotated as "2-hydroxyhepta-2,4-diene-1,7-dioate isomerase/5-carboxymethyl-2-oxohex-3-ene-1,7-dioate decarboxylase". The members of this superfamily include fumarylacetoacetate hydrolase, an enzyme that catalyzes hydrolysis of a carbon−carbon bond in 1,3-dicarbonyl compounds. We postulated that XCC 4068 is 2,4-diketo-3-deoxy-L-fuconate hydrolase that completes the pathway to yield pyruvate and L-lactate, which can be further oxidatively degraded by known enzymes in energy metabolism.

Although we attempted to express and purify all of these proteins with the exception of L-fucose permease, we were successful with only three because of solubility problems: XCC 4069, the putative L-fuconate dehydratase; XCC 4067, the putative 2-keto-3-deoxy-L-fuconate 4-dehydrogenase; and XCC 4070, the putative L-fucopyranoside mutarotase. We frequently encounter solubility problems in our attempts to discover functions of unknown members of the enolase superfamily, with either the superfamily member itself or proteins encoded by the same operon impossible to obtain in soluble form despite efforts involving a survey of vectors and growth conditions (e.g., rich and minimal media, temperature).

XCC 4069 Catalyzes the Dehydration of L-*Fuconate.* Because we were confident that the L-fucose permease function assigned to XCC 4071 would be correct, our first experiment was the incubation of XCC 4069 with L-fuconate. Using ^1H NMR spectroscopy, we observed the dehydration

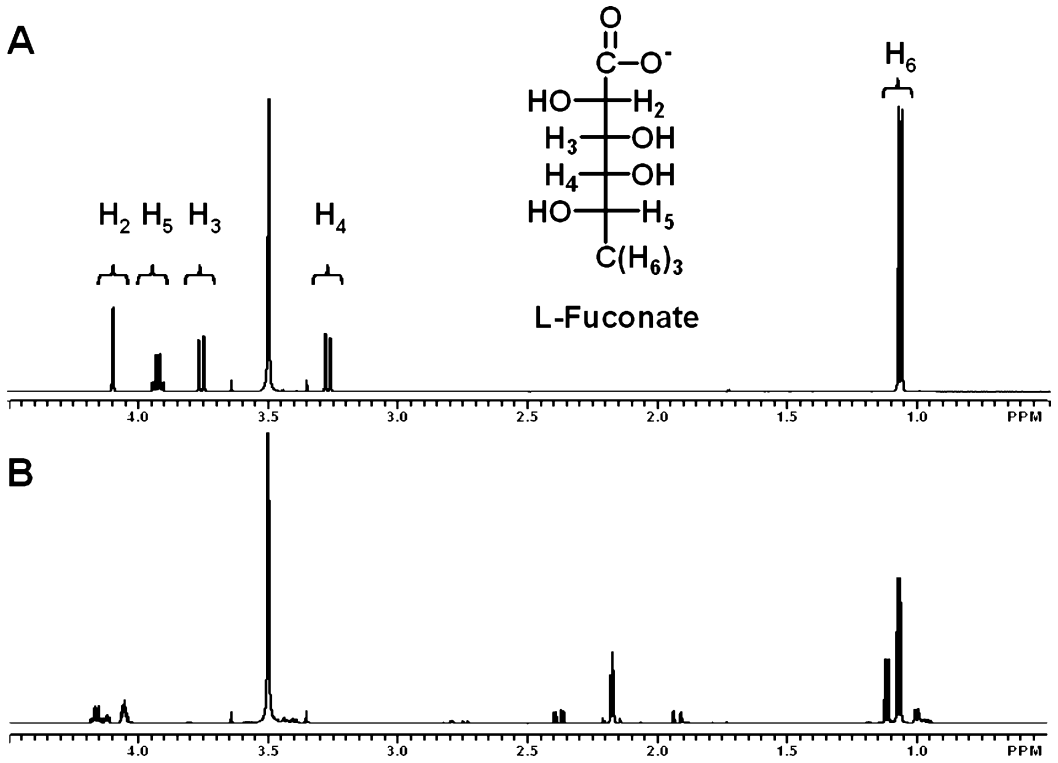


FIGURE 3: (A) ^{1}H NMR spectrum of L-fuconate showing the assignments of all protons. (B) ^{1}H NMR spectrum of 2-keto-3-deoxy-L-fuconate, the dehydrated product of FucD.

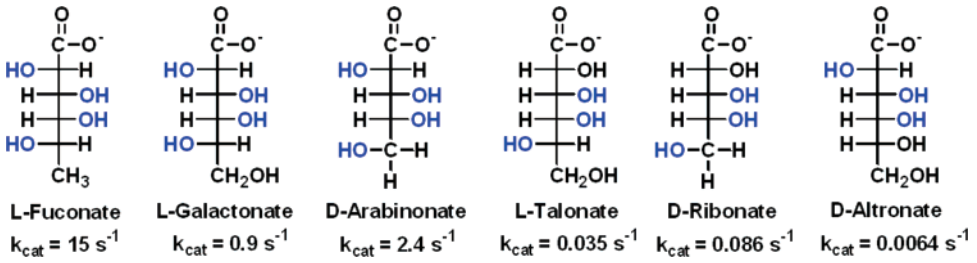


FIGURE 4: Structures of acid sugars within the library that were dehydrated by FucD. The first-order rate constants of dehydration, k_{cat}, for the various acid sugars are listed under the corresponding compounds.

of L-fuconate to 2-keto-3-deoxy-L-fuconate (Figure 3). No reaction was observed with D-fuconate (not shown).

Library Screening of XCC 4069 and Its Functional Assignment as L*-Fuconate Dehydratase*. Although XCC 4069 catalyzes the dehydration of L-fuconate, we sought additional evidence that L-fuconate is the "correct" substrate for this protein. Using a library of 52 mono- and dicarboxylate acid sugars and uronic acids as potential substrates (Figure 1) and screening for dehydration using a semicarbazide end-point measurement, we identified six compounds that were dehydrated by XCC 4069 (Figure 4): L-fuconate, L-galactonate, D-arabinonate, L-talonate, D-ribonate, and D-altronate, although the latter three were not completely converted to product under the conditions of our screening assays (10 mM substrate, 1 μM enzyme, and 16 h at 30 °C).

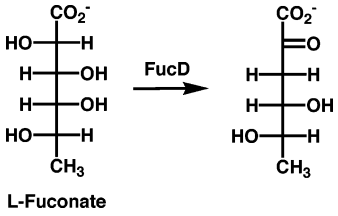
Expecting that the function of XCC 4067 would be 2-keto-3-deoxy-L-fuconate 4-dehydrogenase (vide supra), we determined whether we could use this protein in a coupled-enzyme, spectrophotometric assay for the dehydration reactions catalyzed by XCC 4069. Using the procedure described in the Materials and Methods section, we were able to measure the kinetic constants for the dehydration of L-fuconate, L-galactonate, and D-arabinonate; these are reported in Table 2. L-Fuconate is the preferred substrate, with kinetic constants comparable to those measured for other acid sugar dehydratases of known function in the enolase superfamily. Thus, the library screening persuasively confirms our expectation that L-fuconate is the physiological substrate for XCC 4069. Thus, we assign XCC 4069 as L-fuconate dehydratase (FucD; Scheme 1).

Table 2: Kinetic Parameters of FucD

substrate	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($M^{-1}\,s^{-1}$)
L-fuconate	15 ± 0.2	0.33 ± 0.06	4.5×10^4
L-galactonate	0.9 ± 0.03	6.2 ± 0.4	1.2×10^2
D-arabinonate	2.4 ± 0.07	2.1 ± 0.9	8.3×10^2

We have not isolated and assayed the homologous RTS beta protein found in *H. sapiens* that has been implicated in regulating thymidylate synthase expression (*14*−*16*). However, on the basis of its sequence identity with XCC 4069 (51% identity) and retention of residues that determine substrate specificity (vide infra), we predict that the RTS beta protein also catalyzes the dehydration of L-fuconate. The molecular mechanism by which this reaction may be linked to regulation of thymidylate synthase is unknown,

Scheme 1



L-Fuconate

Table 3: Kinetic Parameters of 2-Keto-3-deoxy-L-fuconate 4-Dehydrogenase

substrate	k_{cat} (s^{-1})	K_M (M)	k_{cat}/K_M ($M^{-1}s^{-1}$)
2-keto-3-deoxy-L-fuconate	6.0 ± 0.2	0.39 ± 0.06	1.5×10^4
2-keto-3-deoxy-L-galactonate	9.8 ± 0.29	6.2 ± 0.4	1.6×10^3
2-keto-3-deoxy-D-arabinonate	7.3 ± 1.1	2.1 ± 0.9	3.5×10^3

Table 4: Rates of Dehydration (k_{cat}) and α-Proton Exchange (k_{exc}) of FucD

substrate	k_{cat} (s^{-1})	k_{exc} (s^{-1})
L-fuconate	15	2.9
L-galactonate	0.9	2.4
D-arabinonate	2.4	4.2
L-talonate	0.035	not detected
D-ribonate	0.086	not detected
D-altronate	0.0064	0.2

although this assignment of the FucD function to the RTS beta protein may facilitate a molecular understanding of that biological process.

L-Fuconate (6-deoxy-L-galactonate) and L-galactonate (6-hydroxy-L-fuconate) share the same configurations at carbon-2 through carbon-5; L-fuconate and D-arabinonate (demethyl-L-fuconate) share the same configurations at carbon-2 through carbon-4, with the 5-OH group of the latter being stereochemically promiscuous due to rotation about the C4−C5 bond. Therefore, on the basis of the observed promiscuity, we conclude that the active site of FucD is specific for acid sugars of common structure, although the efficiency of dehydration depends on the exact molecular structure "remote" from the site of chemistry; as described in a following section, this can be rationalized on the basis of the structures of inhibitor- and substrate-liganded enzymes. Indeed, we expect that our library of acid sugars will not contain the physiological substrates for all of the unknown acid sugar dehydratases in the enolase superfamily. However, we are hopeful that such unknowns sometimes will have promiscuous substrate specificities, such as that observed for FucD, so that the structures of inefficient substrates will provide useful clues regarding the structure/identity of the correct, physiological substrates.

The much slower rates of dehydration of L-talonate, D-ribonate, and D-altronate (Table 4) were estimated from the extent of product formation in the semicarbazide endpoint assays used to screen the library of acid sugars, i.e., at a single concentration of the acid sugar (10 mM). The significance of these slow reactions will be discussed later.

Functional Assignment of XCC 4067 as 2-Keto-3-deoxy-L-fuconate 4-Dehydrogenase. As described in the previous paragraph, we used XCC 4067 as the coupling enzyme to assay FucD, in accord with our expectation that it is 2-keto-3-deoxy-L-fuconate 4-dehydrogenase. We used FucD to generate the dehydration products of L-fuconate, L-galactonate, and D-arabinonate and then used these 2-keto-3-deoxyacids to measure the kinetic constants for XCC 4067; the values are displayed in Table 3. Clearly, the dehydration product obtained from L-fuconate is the preferred substrate, with the value of k_{cat}/K_M in the range expected for the physiological substrate. Thus, we assign XCC 4067 as 2-keto-3-deoxy-L-fuconate 4-dehydrogenase.

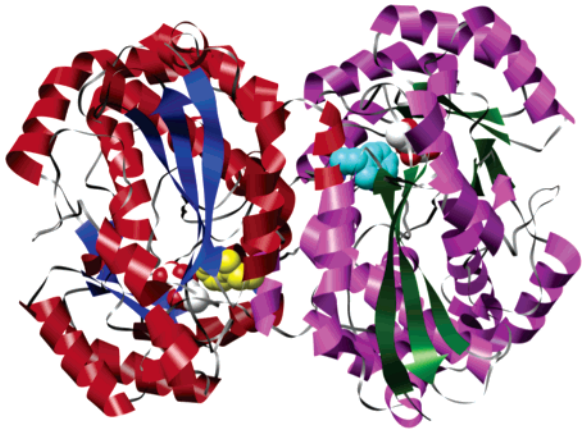


FIGURE 5: The dimer of FucD in the structure of the K220A mutant complexed with L-fuconate (white). The α-helices and β-strands in one polypeptide are colored red and blue, respectively; in the second polypeptide these are colored magenta and green, respectively. Trp 101 from the second polypeptide, colored yellow, participates in the active site of the first polypeptide; Trp 101 from the first polypeptide, colored cyan, participates in the active site of the second polypeptide.

Functional Assignment of XCC 4070 as L-Fucopyranoside Mutarotase. As described in the Materials and Methods section, we used the SD ^{1}H NMR technique (*12*, *13*) and discovered that XCC 4070 accelerates the conversion of L-fucose between its α- and β-pyranose forms. The kinetic constants are $k_{cat} = 2100 \pm 120$ s^{-1}, $K_M = 8.5 \pm 1.7$ mM, and $k_{cat}/K_M = 2.5 \times 10^5$ M^{-1} s^{-1}. Using the same technique, we could not detect any D-fucopyranoside mutarotase activity. We observed that XCC 4070 has measurable mutarotase activity toward L-galactopyranoside and D-arabinopyranoside; however, at 5 mM concentrations, the rates of mutarotation are each reduced ~50-fold from that observed with L-fucopyranoside. Thus, we assign XCC 4070 as L-fucopyranoside mutarotase.

Although we were successful in isolating only three of the six enzymes in the proposed catabolic pathway for L-fucose in Figure 2B, we believe that the functional assignments of these provide persuasive evidence for the occurrence of this pathway in *X. campestris* and other organisms that share orthologues of the proteins encoded by the encoding operon.

Structure of FucD. FucD is a dimer of identical polypeptides, with each polypeptide arranged in two domains: a $(\beta/\alpha)_7\beta$-barrel domain and an $\alpha + \beta$ capping domain that includes both the N- and C-terminal sequences that flank the barrel domain; both share conserved structures with previously characterized members of the superfamily. Although the majority of active site residues is contributed by a single polypeptide chain, Trp 101 from the capping domain of the symmetry-related molecule in the dimer completes the activity site by contributing to the binding pocket for the methyl group of L-fuconate (Figure 5). In the octameric

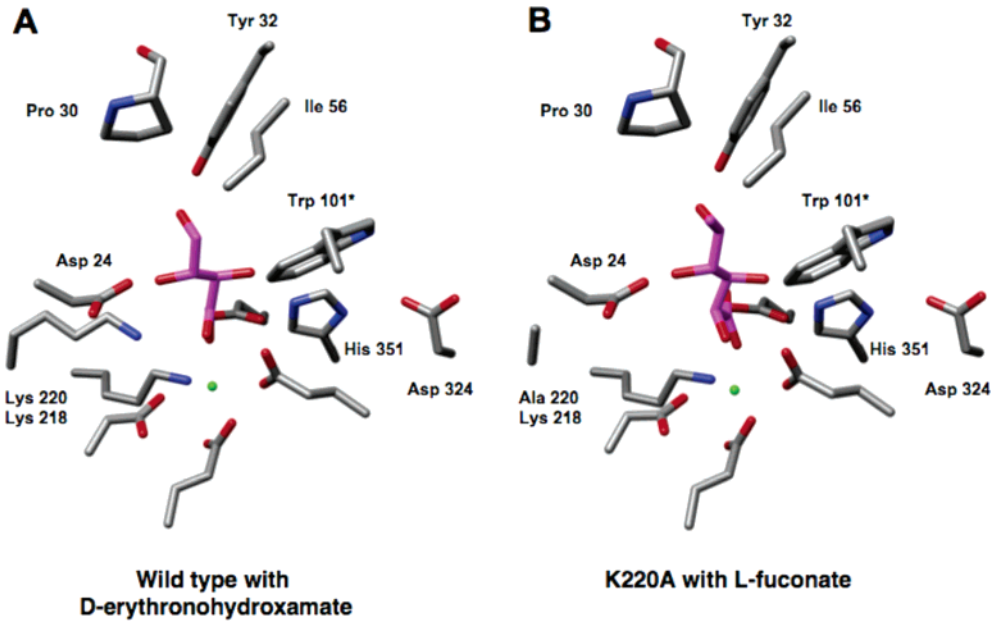


FIGURE 6: Comparison of the active site structures of (A) wild-type FucD complexed with D-erythronohydroxamate and (B) the inactive K220A mutant complexed with L-fuconate.

MR, which is a tetramer of analogous dimers, a Val residue similarly located in its capping domain completes the hydrophobic substrate binding site.

Structures were determined for wild-type FucD both in the absence of a ligand and in the presence of D-erythrono-hydroxamate, an inhibitor and analogue of the enediolate intermediate formed in the dehydration of D-arabinonate (missing the 6-methyl group for synthetic convenience). We also solved the structure of the catalytically inactive K220A mutant in the presence of the L-fuconate substrate (Lys 220 is the base that initiates the reaction by abstraction of the α-proton). Even in the absence of kinetic phenotypes of additional mutants of active site residues, we believe that these structures, in the context of the experiments described in the following sections, allow an accurate description of the mechanism of the FucD-catalyzed reaction.

The active sites of the complexes of wild-type FucD with D-erythronohydroxamate and the inactive mutant K220A mutant with L-fuconate are shown in Figure 6, panels A and B, respectively. Carboxylate ligands for the essential Mg^{2+} are located at the ends of the third, fourth, and fifth β-strands in the barrel domain, Asp 248, Glu 274, and Glu 301, respectively. Both the hydroxamate intermediate analogue and the L-fuconate substrate are bidentate ligands of the Mg^{2+}. The hydroxamate oxygen of the enediolate analogue and the carboxylate oxygen of the substrate that are coordinated to the Mg^{2+} are also hydrogen bonded to the ϵ-ammonium group of Lys 218. Glu 382 located at the end of the eighth β-strand is hydrogen bonded to the second carboxylate oxygen of the substrate. Lys 220 in the wild-type complex with the hydroxamate and His 351 in both complexes, the latter hydrogen bonded to Asp 324, are located on opposite sides of the active site. Lys 220 is positioned to abstract the α-proton from the substrate, and His 351 is hydrogen bonded to the 3-OH group of the substrate/intermediate, consistent with its role as the acid that facilitates the dehydration reaction. As illustrated in Figure 6B, the 3-OH leaving group of L-fuconate is appropriately positioned to be orthogonal to the plane of the enolate anion

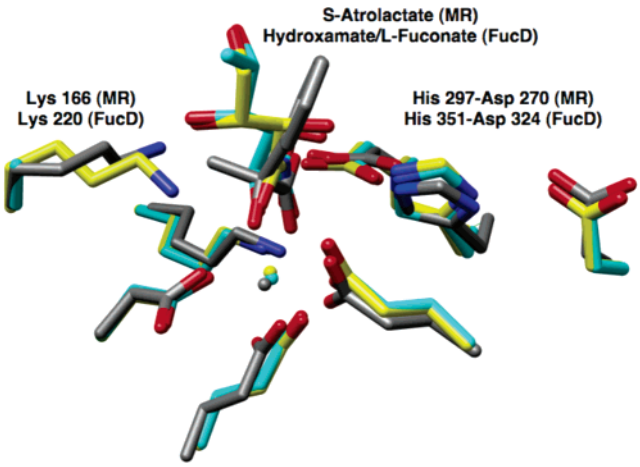


FIGURE 7: Superposition of the active site structures of wild-type MR complexed with (*S*)-atrolactate (gray), wild-type FucD complexed with D-erythronohydroxamate (yellow), and the K220A mutant complexed with L-fuconate (cyan).

intermediate obtained by abstraction of the 2-proton, as required for facile vinylogous elimination.

The liganded structures also reveal the interactions between the active site and the hydroxamate inhibitor/substrate that determine the substrate specificity. In addition to His 351, the 3-OH leaving group is hydrogen bonded to the backbone oxygen of Gly 353. The 4-OH group is hydrogen bonded to Asp 24, and the 5-OH group is hydrogen bonded to both the backbone oxygen of Gly 22 and the phenolic OH group of Tyr 32. The 6-methyl group is located in a hydrophobic pocket formed by Pro 30, Ile 56, and Trp 101 from the symmetry-related polypeptide in the dimer. Gly 22, Asp 24, Pro 30, Tyr 32, and Ile 56 are located in the "20's and 50's" flexible loops in the capping domain; in two polypeptides of the four in the asymmetric unit of the unliganded enzyme, the 20's loop is disordered, with its movement allowing access of the substrate to and egress of the product from the active site.

These interactions are important for understanding the substrate promiscuity discovered by library screening. The

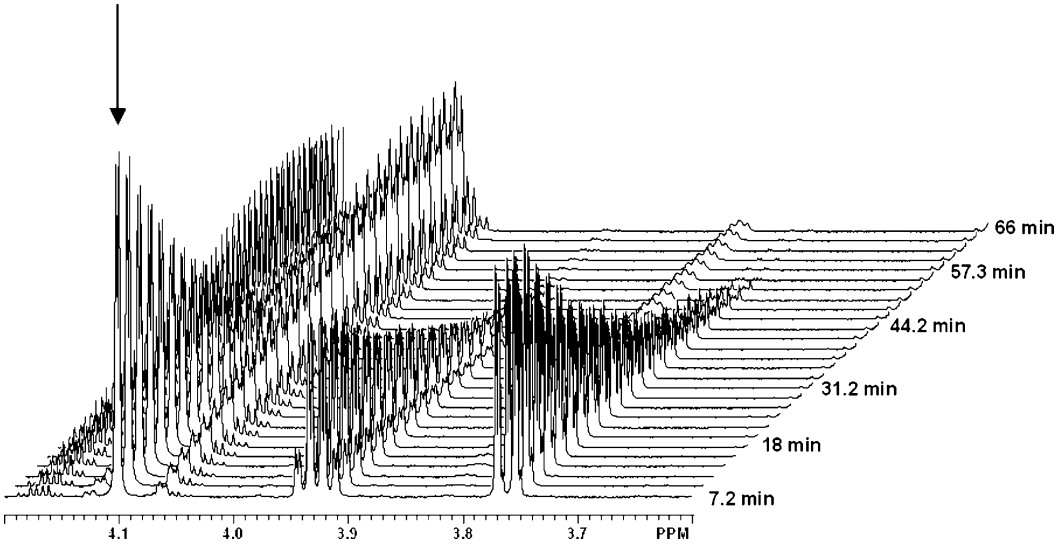


FIGURE 8: ^{1}H NMR spectra of L-fuconate showing the FucD-catalyzed exchange of the α-proton with solvent-derived deuterium with time. The diminishing α-proton peak is indicated by the arrow.

"best" substrates for dehydration, L-fuconate, L-galactonate, and D-arabinonate, share stereochemically equivalent hydroxyl groups at carbon-3 through carbon-5 that are able to participate in analogous hydrogen-bonding interactions (the 5-hydroxymethyl group of D-arabinonate is stereochemically promiscuous). The substrates for epimerization/dehydration, L-talonate and D-ribonate that are 2-epimers of L-galactonate and D-arabinonate, also share stereochemically equivalent hydroxyl groups at carbon-3 through carbon-5, thereby allowing these to bind in the active site, although the 2-proton would necessarily be presented to His 351 for abstraction instead of Lys 220. In contrast, D-altronate that undergoes "rapid" exchange but "slow" dehydration has the opposite configuration at carbon-5. Presumably, this configurational change prevents equivalent hydrogen bonding with the backbone oxygen of Gly 22 and the phenolic OH group of Tyr 32, so the conformation of the substrate is altered such that the 3-OH group is not correctly positioned for acid-catalyzed *β*-elimination.

A superposition of the active sites of wild-type FucD complexed with D-erythronohydroxamate, the inactive K220A mutant complexed with L-fuconate, and wild-type MR complexed with the inhibitor (*S*)-atrolactate [(*S*)-α-methylmandelate] is presented in Figure 7. Inspection of this figure reveals that both the Lys located at the end of the second *β*-strand (Lys 220 in FucD and Lys 166 in MR) and the His-Asp dyads located at the ends of the seventh and sixth *β*-strands (His 351 and Asp 324 in FucD and His 297 and Asp 270 in MR) are positioned equivalently in the two active sites. Thus, the position of His 351 in the active site of FucD is appropriate for catalysis of either departure of the 3-OH leaving group from L-fuconate or abstraction of the 2-proton from either L-talonate and D-ribonate, thereby explaining the observed ability of FucD to catalyze both dehydration and epimerization reactions (vide infra). The rates of the dehydration and epimerization reactions differ, however, likely as the result of subtle kinetically important changes in the disposition of the 2-epimeric L-talonate and D-ribonate substrates with respect to His 351.

Mechanism of the FucD-Catalyzed Reaction: Exchange of the α-Proton in Competition with Dehydration. The time course of dehydration of L-fuconate catalyzed by FucD was monitored using ^{1}H NMR spectroscopy. The intensities of the protons associated with L-fuconate decreased with time, concomitant with the appearance of those associated with the keto tautomer of the dehydration product, 2-keto-3-deoxy-L-fuconate (Figure 8). However, upon closer inspection of the spectra, the intensity of the resonance associated with the α-proton (on C-2) of L-fuconate decreased more rapidly than those of the remaining protons of the substrate. As a result of this exchange reaction with solvent deuterium, the resonance associated with the proton on C-3 (a doublet of doublets) collapsed to a doublet as the reaction proceeded (Figures 8 and 9A). This exchange reaction, i.e., formation of L-[2-^{2}H]-fuconate, can be explained by the abstraction of the α-proton of L-fuconate by a polyprotic active site base, e.g., Lys 220, to generate the Mg^{2+}-stablilized enediolate intermediate; this intermediate partitions between (1) vinylogous *β*-elimination to accomplish dehydration of the enediolate intermediate to yield the enol tautomer of the dehydration product and (2) deuteriation by the conjugate acid of the polyprotic base to yield L-[2-^{2}H]-fuconate (Scheme 2).

This exchange reaction is reminiscent of that observed in the MR-catalyzed reaction when (*S*)-mandelate, but not (*R*)-mandelate, is used as substrate (*20*). In the active site of MR, one carboxylate oxygen and the 2-OH group of either (*S*)- or (*R*)-mandelate are coordinated to the essential Mg^{2+}, so the α-hydrogen of (*S*)-mandelate is "presented" to Lys 166 located at the end of the second *β*-strand; i.e., Lys 166 is the *S*-specific base. Because L-fuconate binds in the active site of FucD with an analogous coordination geometry, the 2-proton is presented to Lys 220 at the end of the second *β*-strand. Thus, the exchange reaction observed with L-fuconate is persuasive evidence that Lys 220 initiates the reaction by abstraction of the α-proton.

We also examined the ability of FucD to catalyze exchange of the α-proton of the remaining members of our acid sugar library. Exchange was observed for only three additional acid sugars, L-galactonate, D-arabinonate, and D-altronate, all three of which are also substrates for dehydration (panels B, C, and D of Figure 9, respectively); the values of the exchange

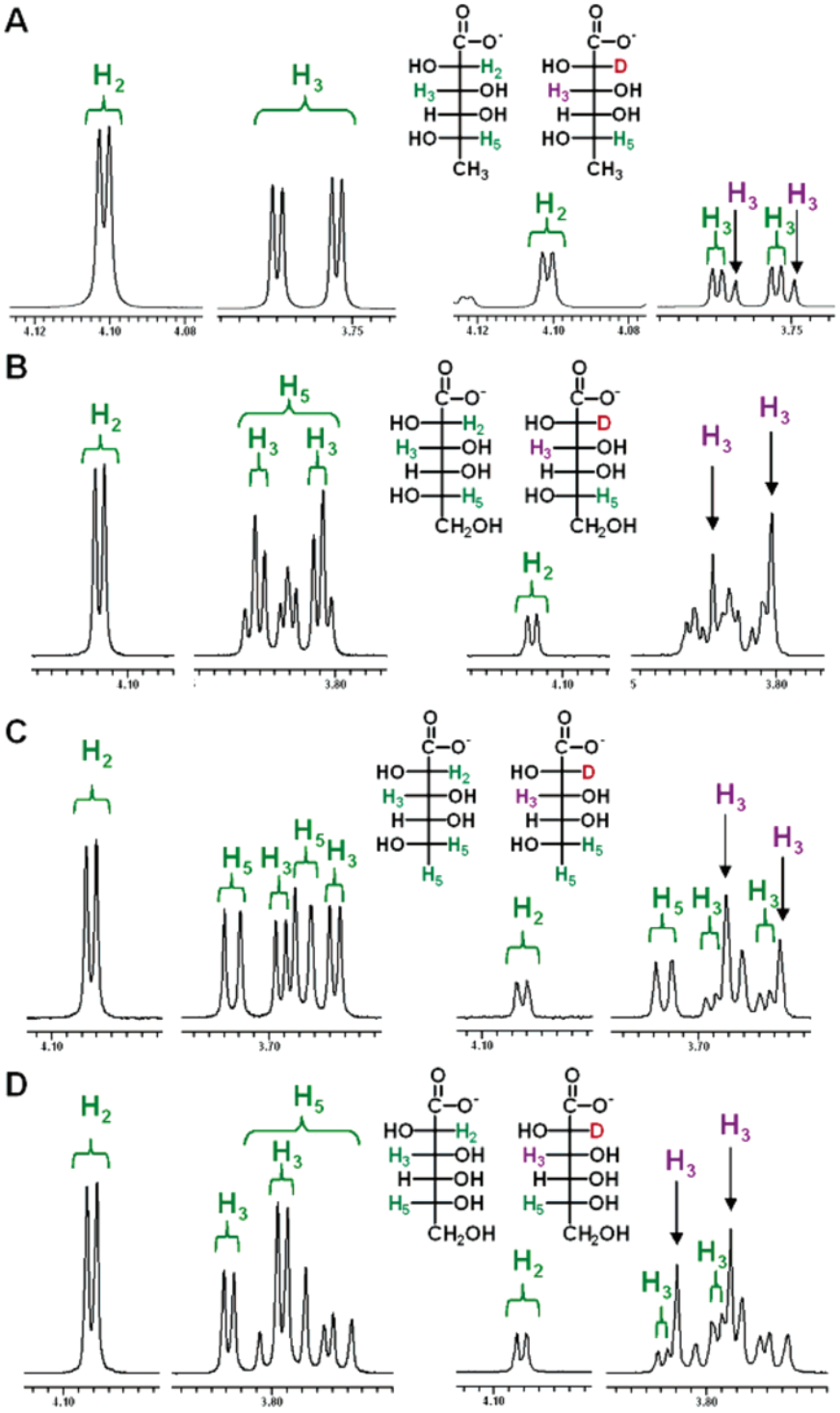


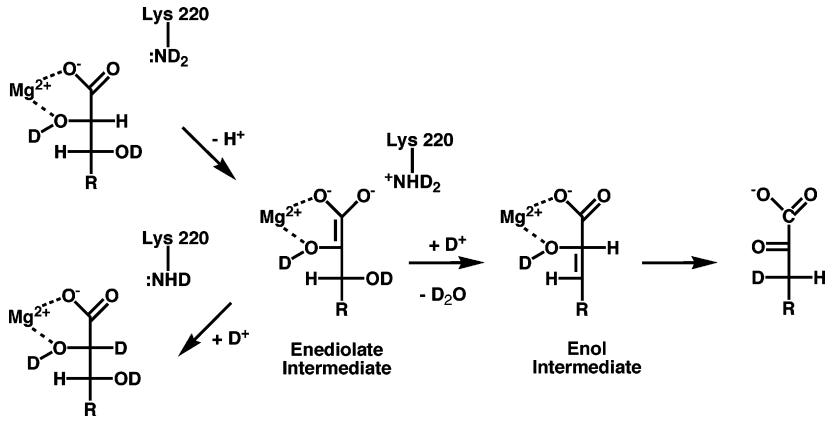
FIGURE 9: Representative partial ^{1}H NMR spectrum of (A) L-fuconate, (B) L-galactonate, (C) D-arabinonate, and (D) D-altronate after incubation with FucD in D_2O buffer. The doublet peaks associated with the proton of C3 of L-[2(*S*)-^{2}H]substrate are indicated by magenta H_3.

rates measured as described in the Materials and Methods section are compared with those for dehydration in Table 3. Exchange was not observed for either L-talonate or D-ribonate, although these are "better" substrates for dehydration than D-altronate.

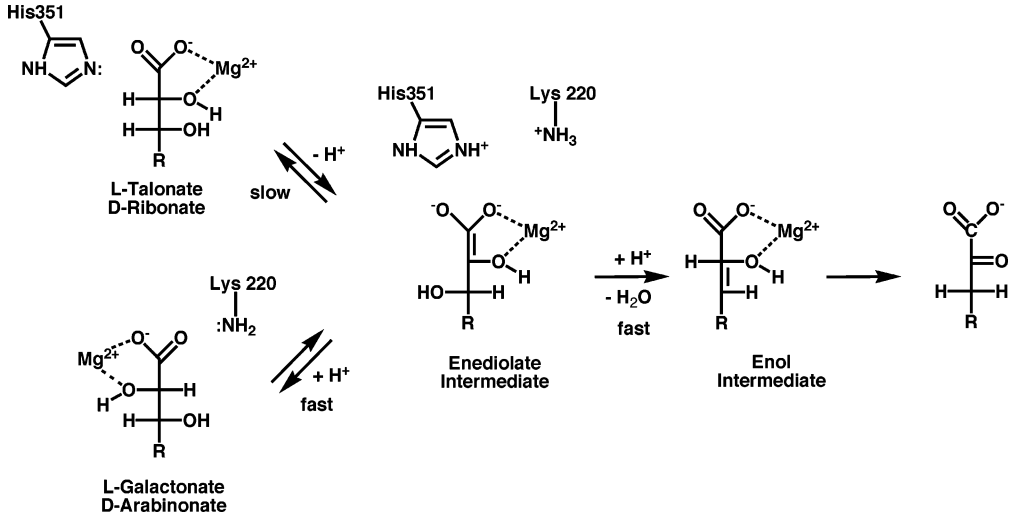
As noted previously, L-fuconate and L-galactonate share the same configurations at carbon-2 through carbon-5; L-fuconate and D-arabinonate share the same configurations at carbon-2 through carbon-4, with the latter configurationally promiscuous at carbon-5. However, in contrast to the relative rates of the dehydration and the exchange reactions measured for L-fuconate, the rate constants for exchange of the 2-protons of L-galactonate and D-arabinonate exceed those for dehydration, highlighting the importance of the structure of the distal portion of the substrate in positioning the enediolate intermediate for efficient vinylogous β-elimination.

The configuration of carbon-5 of D-altronate differs from those in L-fuconate and L-galactonate. Although the rate of exchange of the 2-proton of D-altronate is ~10-fold less than those measured for L-fuconate, L-galactonate, and D-arabinonate, the rate of dehydration is 1000-fold less. Again,

Scheme 2



Scheme 3



this difference highlights the importance of the distal portion of the substrate in determining reactivity for dehydration, presumably as the result of subtle changes in active site geometry. Although D-altronate dehydratase is not yet a known physiological function of any member of the enolase superfamily (the D-altronate dehydratase in D-galacturonate catabolism is a member of a distinct Fe^{2+}-dependent superfamily), the propensity of D-altronate for exchange of the α-proton but not dehydration in the active site of FucD suggests that in vivo and/or in vitro evolution of this function is possible.

Mechanism of the FucD-Catalyzed Reaction: Epimerization in Competition with Dehydration. Although both L-talonate and D-ribonate are slow substrates for dehydration, we did not detect any exchange of the α-proton of either with solvent. On the basis of the established structure−function relationships for MR and the liganded structures of FucD, the α-protons of L-talonate and D-ribonate should be presented to His 351 located at end of the seventh β-strand. Because the conjugate acid of His is not torsiosymmetric, in contrast to the conjugate acid of Lys 220, exchange of the α-protons of L-talonate and D-ribonate with solvent hydrogen is not expected. Indeed, the α-proton of (*R*)-mandelate is abstracted by His 297, and no exchange of the α-proton is observed in competition with 1,1-proton transfer (*21*).

However, if His 351 were able to abstract the α-proton of either L-talonate or D-ribonate to generate the same enediolate intermediate obtained from L-galactonate or D-arabinonate, respectively, the conjugate acid of Lys 220 might be able to protonate this intermediate on the opposite to generate L-galactonate or D-arabinonate, respectively, in competition with vinylogous elimination. As illustrated in Scheme 3, the ability to observe epimerization is the result of the magnitudes of the energy barriers for partitioning of the enediolate intermediate among irreversible dehydration, protonation on the same face to return to substrate, and protonation on the opposite face to give the epimerized product.

When L-galactonate or D-arabinonate was incubated with FucD in H_2O-containing buffer, we were unable to detect the L-talonate or D-arabinonate, respectively (data not shown), presumably because the rate of vinylogous β-elimination is much greater than protonation of the opposite face to give L-talonate or D-ribonate. However, in the reverse direction, when L-talonate or D-ribonate was substrate, we were able to detect small amounts of the epimerized L-galactonate or D-arabinonate by the observation of the C2 α-protons at the characteristic chemical shifts of 4.12 and 4.09 ppm, respectively (Figure 10). In this case, when the enediolate intermediate is formed from L-talonate or D-ribonate (the same enediolate intermediate formed from L-galactonate or D-arabinonate), albeit slowly by abstraction of the proton on

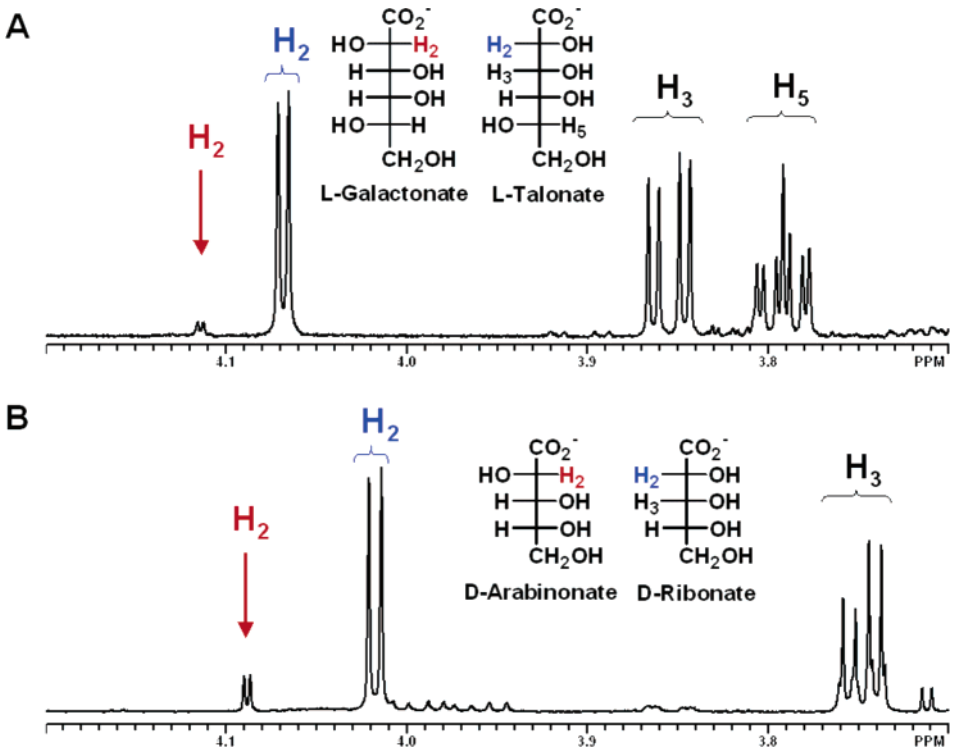


FIGURE 10: (A) Partial ^{1}H NMR spectrum of L-talonate incubated with FucD showing the presence of the α-proton of the epimerized product L-galactonate. (B) Partial ^{1}H NMR spectrum of D-ribonate incubated with FucD showing the presence of the α-proton of the epimerized product D-arabinonate.

carbon-2 by His 351, it can partition between vinylogous *β*-elimination and protonation on the opposite face to give L-galactonate or D-arabinonate (the reaction manifolds accessible to the enediolate intermediate formed from L-galactonate and D-arabinonate).

That the FucD catalyzes the epimerization of L-talonate and D-ribonate, albeit at slow rates, reveals a functional promiscuity that, in retrospect, is not surprising given the similarity of the active sites of FucD and MR (Figure 7). Indeed, in unpublished studies using our acid sugar library, we have discovered a homologue of MR and FucD that catalyzes the dehydration of L-talarate and galactarate as well as their epimerization (in species of *Salmonella* and several other bacteria; W. S. Yew and J. A. Gerlt, unpublished observations). Sequence alignments allow the prediction that the active site of L-talarate/galactarate dehydratase/epimerase shares the KxK motifs and the His-Asp dyads found in the active sites of FucD and MR. Although a structure is not yet available for this enzyme, its structure may reveal the structural differences in the active site that determine the relative efficiencies of acid sugar dehydration and epimerization reactions and, therefore, direct the evolution of mechanisms in the MR subgroup of the enolase superfamily. In the MLE subgroup, we have demonstrated that changes in the substrate specificity determinants are sufficient to interconvert dehydration (*o*-succinylbenzoate synthase) and 1,1-proton transfer (L-Ala-D/L-Glu epimerase) functions (*22*, *23*).

Stereochemical Course of the FucD-Catalyzed Dehydration Reaction. We also examined the stereochemical course of the replacement of the 3-OH group of L-fuconate with solvent-derived hydrogen so that we could determine the relative orientation of the leaving group and the general acid that catalyzes ketonization of the enol product formed by vinylogous *β*-elimination of water to yield the 2-keto-3-deoxy product.

The product exists as a mixture of α- and *β*-furanosyl hemiketals (Figure 11A) as assessed by both ^{1}H and ^{13}C NMR spectroscopies. From the magnitudes of the vicinal C3−C4 ^{1}H−^{1}H coupling constants as well as NOE measurements, we determined that the larger coupling constant is associated with the 3-*pro-S* hydrogen. When L-fuconate is dehydrated in D_2O-containing buffer, one prochiral hydrogen of C3 of each anomer was stereospecifically deuteriated in the 3-*pro-S* position to yield 2-keto-3-deoxy-[3(*S*)-^{2}H]-L-fuconate (Figure 11B). Thus, ketonization of the enol intermediate derived from vinylogous *β*-elimination is enzyme-catalyzed, and the departing 3-OH group is replaced with solvent deuterium with inversion of configuration. We performed a similar analysis with L-galactonate and again observed that the 3-OH group is replaced with solvent deuterium with inversion of configuration (data not shown). With the expectation that dehydration proceeds via an anti-stereochemical course, given the structures of the liganded complexes reported in this report, the solvent-derived hydrogen is delivered from the side of the active site on which Lys 220 is located.

The replacement of the 3-OH group by solvent-derived hydrogen with inversion of configuration contrasts with the retention of configuration previously reported for the GalD- and GlucD-catalyzed reactions (*24*). In the GalD-catalyzed reaction, a His located at the end of the third *β*-strand is the acid that facilitates the departure of the 3-OH group in an anti-elimination reaction (*24*); in the GlucD-catalyzed reaction, the His of the His-Asp dyad is the single acid/base catalyst that initiates the dehydration reaction by abstraction of the α-proton and facilitates the departure of the 3-OH group in a syn-elimination reaction (*25*). Thus, no structure-

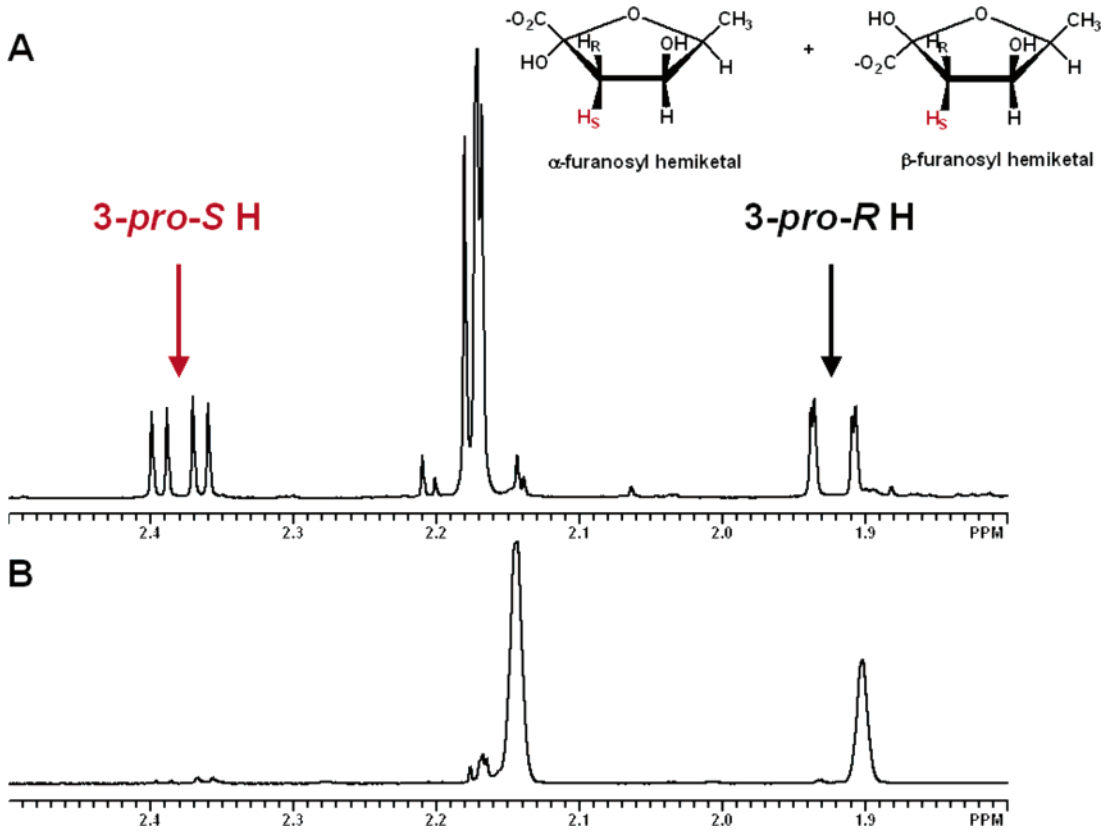


FIGURE 11: Partial ^{1}H NMR spectrum of 2-keto-3-deoxy-L-fuconate (A) obtained by dehydration of L-fuconate in H_2O, showing the resonances associated with the 3-*pro*-*S* and 3-*pro*-*R* protons, and (B) obtained by dehyration of L-fuconate in D_2O buffer.

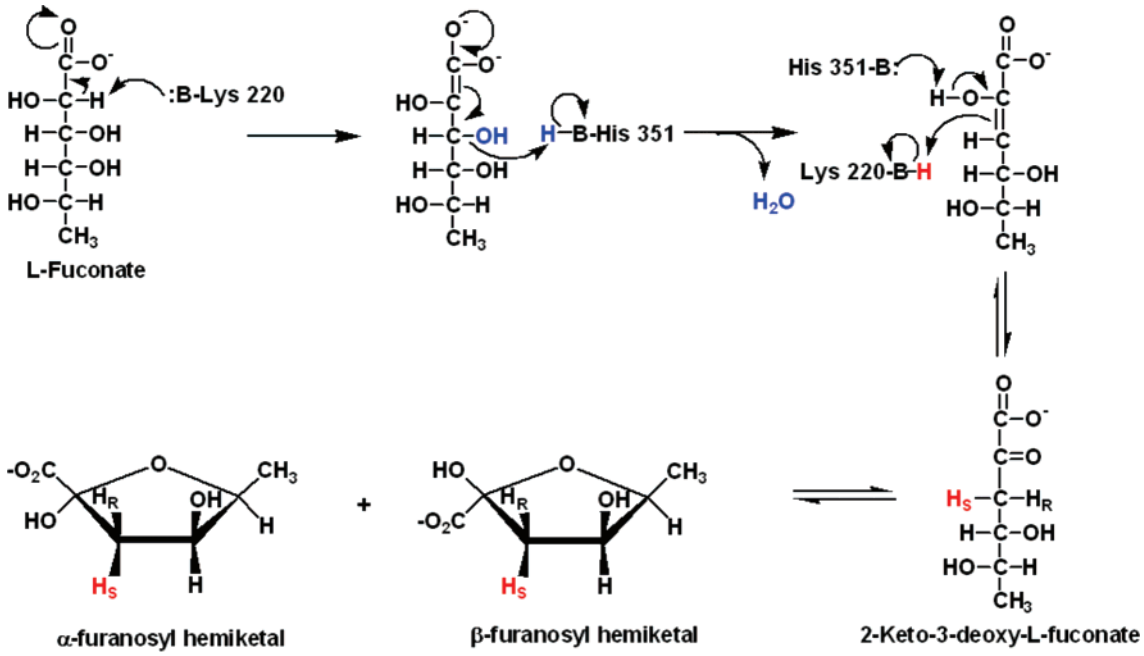


FIGURE 12: Proposed mechanism of the FucD-catalyzed reaction.

based strategy is conserved for catalyzing the ketonization of the enol intermediates obtained from stereochemically diverse dehydration reactions.

Conclusions. We used both operon context and a library of acid sugars to independently assign the FucD function to a member of the MR subgroup encoded by the *X. campestris pv. campestris str.* ATCC 33913 genome. That both approaches yield the same function validates the use of a screening strategy to assign functions to other unknown members of the enolase superfamily.

On the basis of structural and functional evidence, the dehydration of L-fuconate catalyzed by FucD is initiated by abstraction of the 2-proton by Lys 220 to generate a transiently stable enediolate intermediate that partitions

between vinylogous β-elimination and reprotonation to reform the substrate (Figure 12). The conjugate acid of His 351 in the His 351-Asp 324 hydrogen-bonded dyad is positioned to function as the acid catalyst in the β-elimination reaction. The ketonization of the resulting enol intermediate is enzyme-catalyzed with inversion of configuration at carbon-3, suggesting the participation of the conjugate acid of Lys 220 in the final partial reaction.

FucD is able to catalyze the dehydration of structural analogues of L-fuconate, including L-galactonate, D-arabinonate, and D-altronate that share the same configuration at the reactive 2- and 3-carbons as well as carbon-4. Although the dehydration of D-altronate is very slow, the rate of abstraction of its α-proton, as assessed by exchange with solvent deuterium, is fast, thereby revealing the importance of the configuration of carbon-5 on the mechanism of the reaction.

Finally, FucD also catalyzes the slow dehydration and epimerization of L-talonate and D-ribonate that are epimers of the physiological substrate at carbon-2. The ability of FucD to catalyze these epimerization reactions is explained by an active site structure that closely resembles that of MR, although the physiological role of the conjugate acid of the His at the end of the seventh β-strand in FucD is catalysis of the β-elimination of the 3-OH group rather than proton abstraction from/protonation of C-2 by the homologous His in the 1,1-proton transfer reaction catalyzed by MR. The functional promiscuity discovered for FucD highlights possible mechanisms for evolution of function in the enolase superfamily.

SUPPORTING INFORMATION AVAILABLE

Descriptions of the syntheses of the mono- and diacid sugars in the library used for substrate screening. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Neidhart, D. J., Kenyon, G. L., Gerlt, J. A., and Petsko, G. A. (1990) Mandelate racemase and muconate lactonizing enzyme are mechanistically distinct and structurally homologous, *Nature 347*, 692−694.
2. 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.
3. 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.
4. 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*, 4455−4462.
5. Gerlt, J. A., and Raushel, F. M. (2003) Evolution of function in (beta/alpha)(8)-barrel enzymes, *Curr. Opin. Chem. Biol. 7*, 252−264.
6. Babbitt, P. C., Hasson, M. S., Wedekind, J. E., Palmer, D. R., Barrett, W. C., Reed, G. H., Rayment, I., Ringe, D., Kenyon, G. L., and Gerlt, J. A. (1996) The enolase superfamily: a general strategy for enzyme-catalyzed abstraction of the alpha-protons of carboxylic acids, *Biochemistry 35*, 16489−16501.
7. Kalyanaraman, C., Bernacki, K., and Jacobson, M. P. (2005) Virtual screening against highly charged active sites: identifying substrates of alpha-beta barrel enzymes, *Biochemistry 44*, 2059−2071.
8. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, in *Methods in Enzymology* (Carter, C. W. J., Sweet, R. M., Abelson, J. N., and Simon, M. I., Eds.) pp 307−326, Academic Press, New York.
9. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 849−861.
10. 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*.
11. Jones, A. T. (1985) Interactive computer graphics: FRODO, *Methods Enzymol. 115*, 157−171.
12. Ryu, K. S., Kim, C., Park, C., and Choi, B. S. (2004) NMR analysis of enzyme-catalyzed and free-equilibrium mutarotation kinetics of monosaccharides, *J. Am. Chem. Soc. 126*, 9180−9181.
13. Ryu, K. S., Kim, C., Kim, I., Yoo, S., Choi, B. S., and Park, C. (2004) NMR application probes a novel and ubiquitous family of enzymes that alter monosaccharide configuration, *J. Biol. Chem. 279*, 25544−25548.
14. Dolnick, B. J., Lu, K., Yin, M. B., and Rustum, Y. M. (1997) Recent advances in the study of rTS proteins. rTS expression during growth and in response to thymidylate synthase inhibitors in human tumor cells, *Adv. Enzyme Regul. 37*, 95−109.
15. Dolnick, B. J., Angelino, N. J., Dolnick, R., and Sufrin, J. R. (2003) A novel function for the rTS gene, *Cancer Biol. Ther. 2*, 364−369.
16. Dolnick, B. J. (2005) The rTS signaling pathway as a target for drug development, *Clin. Colorectal Cancer 5*, 57−60.
17. Gunn, F. J., Tate, C. G., and Henderson, P. J. (1994) Identification of a novel sugar-H^+ symport protein, FucP, for transport of L-fucose into *Escherichia coli*, *Mol. Microbiol. 12*, 799−809.
18. Lu, Z., and Lin, E. C. (1989) The nucleotide sequence of *Escherichia coli* genes for L-fucose dissimilation, *Nucleic Acids Res. 17*, 4883−4884.
19. Ryu, K. S., Kim, J. I., Cho, S. J., Park, D., Park, C., Cheong, H. K., Lee, J. O., and Choi, B. S. (2005) Structural insights into the monosaccharide specificity of *Escherichia coli* rhamnose mutarotase, *J. Mol. Biol. 349*, 153−162.
20. 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.
21. Powers, V. M., Koo, C. W., Kenyon, G. L., Gerlt, J. A., and Kozarich, J. W. (1991) Mechanism of the reaction catalyzed by mandelate racemase. 1. Chemical and kinetic evidence for a two-base mechanism, *Biochemistry 30*, 9255−9263.
22. Schmidt, D. M., Mundorff, E. C., Dojka, M., Bermudez, E., Ness, J. E., Govindarajan, S., Babbitt, P. C., Minshull, J., and Gerlt, J. A. (2003) Evolutionary potential of (beta/alpha)8-barrels: functional promiscuity produced by single substitutions in the enolase superfamily, *Biochemistry 42*, 8387−8393.
23. Vick, J. E., Schmidt, D. M., and Gerlt, J. A. (2005) Evolutionary potential of (beta/alpha)(8)-barrels: In vitro enhancement of a "new" reaction in the enolase superfamily, *Biochemistry 44*, 11722−11729.
24. 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.
25. 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.

BI061687O