Evolution of Enzymatic Activities in the Enolase Superfamily: Identification of the General Acid Catalyst in theActive Site of D-Glucarate Dehydratase from *Escherichia coli*[†,‡]

Andrew M. Gulick,[§,||] Brian K. Hubbard,[||,⊥,#] John A. Gerlt,*[,⊥] and Ivan Rayment*[,§]

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

Received April 11, 2001; Revised Manuscript Received June 15, 2001

ABSTRACT: D-Glucarate dehydratase from *Escherichia coli* (GlucD), a member of the enolase superfamily, catalyzes the dehydration of both D-glucarate and L-idarate to form 5-keto-4-deoxy-D-glucarate (KDG). Previous mutagenesis and structural studies identified Lys 207 and the His 339−Asp 313 dyad as the general basic catalysts that abstract the C5 proton from L-idarate and D-glucarate, respectively, thereby initiating the reaction by formation of a stabilized enediolate anion intermediate [Gulick, A. M., Hubbard, B. K., Gerlt, J. A., and Rayment, I. (2000) *Biochemistry 39,* 4590−4602]. The vinylogous elimination of the 4-OH group from this intermediate presumably requires a general acid catalyst. The structure of GlucD with KDG and 4-deoxy-D-glucarate bound in the active site revealed that only His 339 and Asn 341 are proximal to the presumed position of the 4-OH leaving group. The N341D and N341L mutants of GlucD were constructed and subjected to both mechanistic and structural analyses. The N341L but not N341D mutant catalyzed the dehydrofluorination of 4-deoxy-4-fluoro-D-glucarate, demonstrating that in this mutant the initial proton abstraction from C5 can be decoupled from elimination of the leaving group from C4. The kinetic properties and structures of these mutants suggest that either Asn 341 participates in catalysis as the general acid that facilitates the departure of the 4-leaving group or is essential for proper positioning of His 339. In the latter scenario, His 339 would function not only as the general base that abstracts the C5 proton from D-glucarate but also as the general acid that catalyzes both the departure of the 4-OH group and the stereospecific incorporation of solvent hydrogen with retention of configuration to form the KDG product. The involvement of a single functional group in this reaction highlights the plasticity of the active site design in members of the enolase superfamily.

We (*1*−*5*) and others (*6*, *7*) have described a number of mechanistically diverse superfamilies whose members catalyze different overall reactions but use a conserved structural framework to stabilize a common intermediate in the reaction mechanism. The enolase superfamily is the most well-characterized superfamily whose members catalyze different overall reactions. The members of this superfamily have a conserved bidomain structure: a mixed α/β domain formed from the N- and C-termini and a larger $(\beta/\alpha)_7\beta$-barrel domain (*8*−*15*). The barrel domain contains the functional groups that are directly involved in catalysis.

The reactions catalyzed by members of the enolase superfamily are initiated by abstraction of a proton adjacent to a carboxylate anion to generate an enolate anion intermediate (*1*). The enzymes contain one or two general acid/base catalysts on opposite faces of the barrel domain that are responsible for abstraction of the α-proton from the substrate: a Lys at the end of the second β-strand and either a Lys at the end of the sixth β-strand or a His-Asp dyad at the ends of the sixth (Asp) and seventh (His) β-strands. Stabilization of the enolate anion intermediate occurs by its interaction with a divalent metal ion (usually Mg^{2+}) that is coordinated by three (usually anionic) functional groups located at the ends of the third, fourth, and fifth β-strands of the barrel domain. After this initial shared chemical step, the reactions diverge among different family members reflecting alternate fates of the enolic intermediate. This is possible because the barrel domain is well-designed for providing the architecture for the divergent reactions: the functional groups are located on separate β-strands so their positions and identities can be independently varied in the divergent evolution of new catalysts.

[†] This research was supported by Grants GM-40570 (to J.A.G.), GM-52594 (to J.A.G. and I.R.), and AR-35186 (to I.R.) from the National Institutes of Health. Use of the Argonne National Laboratory Structural Biology Center beamline at the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Energy Research, under Contract W-31-109-ENG-38.

[‡] The X-ray coordinates of the different GlucD complexes have been deposited in the Protein Data (1JDF, N341D Mutant; 1JCT, N341L Mutant).

* To whom correspondence should be addressed. (I.R.) Phone: (608) 262-0437. Fax: (608) 262-1319. E-mail: ivan_rayment@biochem.wisc.edu. (J.A.G.) Phone: (217) 244-7414. Fax: (217) 265-0385. E-mail: j-gerlt@uiuc.edu.

[§] University of Wisconsin.

[||] These two authors contributed equally to this work.

[⊥] University of Illinois.

[#] Current address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School-LHRRB, 45 Shattuck Street, Boston, MA 02115.

10.1021/bi010733b CCC: $20.00
Published on Web 08/01/2001

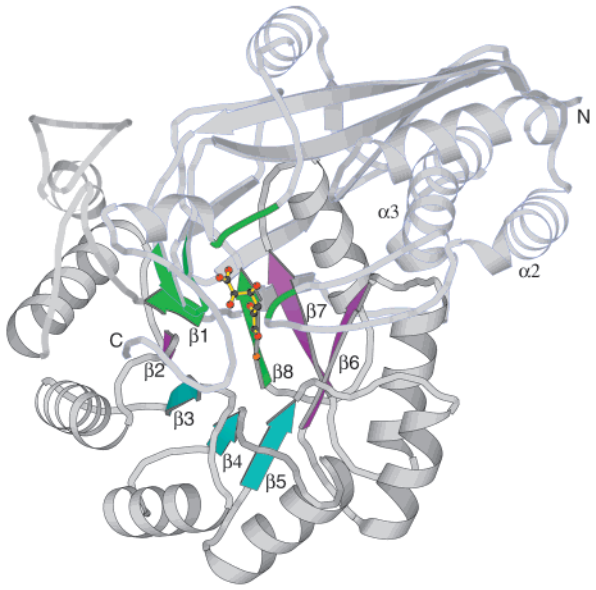
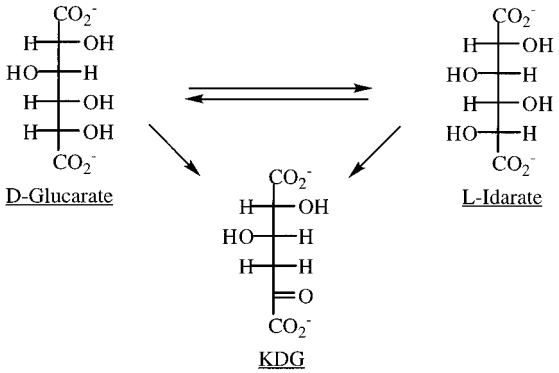


FIGURE 1: Ribbon diagram of GlucD. A single subunit of the GlucD protein is illustrated. The figure shows a subunit of the N341D mutant protein with 5KDG bound in the active site (PDB accession number 1JDF). The two-domain nature of the protein is apparent, with the capping domain composed of the N-terminal 145 residues and the C-terminal 80 residues, and the barrel domain in the center of the sequence. The strands of the barrel are labeled $\beta1-\beta8$. The active site is illustrated by the presence of a complex of Mg^{2+}-5KDG shown in ball-and-stick. The regions of the protein that are involved at the active site are colored according to the function of residues contributed to the active site: the three strands that donate the metal binding ligands are blue; the strands that donate the catalytic bases are magenta; and the two strands, as well as two regions from the capping domain that are involved in substrate specificity are shown in green.

As an example of this design, the structurally independent nature of the functional groups for the D-glucarate dehydratase (GlucD)[1] from *Escherichia coli* is illustrated in Figure



1. The three metal binding ligands, Asp 235, Glu 260, and Asn 289, are located on barrel strands $\beta3$, $\beta4$, and $\beta5$. The catalytic bases, Lys 207 and the His 339−Asp 313 dyad, are located on strands $\beta2$ and $\beta7/\beta6$, respectively. Four functional groups that determine substrate specificity also are located on separate structural elements: Tyr 150 on strand $\beta1$, His 368 on strand $\beta7$, Thr 103 on the loop between α2 and α3 in the N-terminal domain, and Arg 422 on the C-terminal loop.

Seven members of the enolase superfamily have been structurally characterized (*8*−*15*). The first three members for which high-resolution structures were determined, mandelate racemase (*8*), muconate lactonizing enzyme (*10*, *11*), and enolase (*9*), serve as the structural paradigms for the three subfamilies into which all subsequently recognized members can be placed on the basis of sequence homology. The more recently characterized members of this superfamily are the D-glucarate dehydratases (GlucD) from both *Pseudomonas putida* (*12*) and *E. coli* (*15*), D-galactonate dehydratase (GalD) from *E. coli* (*13*), and *o*-succinylbenzoate synthase from *E. coli* (OSBS) (*14*). The structures of five of these, MR (*16*), enolase (*17*), GlucD (*15*), GalD (*13*), and OSBS (*14*), have been determined in the presence of substrates or substrate analogues. This structural information, together with biochemical characterization of the enzymes and appropriate active site mutants, has provided considerable insight into the relationships between active site structure and mechanism. In particular, these studies have identified many of the amino acids that function as critical acids and bases during catalysis which is important for understanding the evolutionary relationships between these enzymes.

For the reactions catalyzed by MR, MLE, and OSBS, the reaction mechanisms can be explained using only the general acid/base catalysts located at the ends of the second (Lys-X-Lys), sixth (Lys), and sixth/seventh (His-Asp dyad) β-strands. However, the dehydration reactions catalyzed by enolase, GlucD, and GalD might be expected to involve an additional general acid catalyst to facilitate the vinylogous elimination of the β-OH group from the enolate anion intermediate. In the case of enolase, Glu 211 that is located on an extended loop between the second and third β-strands of the barrel domain serves as the acid catalyst (*17*). In the case of GalD, His 185 located at the end of the third β-strand has been identified as this additional catalyst (*13*). However, the general acid catalyst in the GlucD-catalyzed reaction had not been identified, thereby restricting a full understanding of structure/function relationships in this active site.

GlucD catalyzes the dehydration of both D-glucarate and L-idarate to form 5-keto-4-deoxyglucarate (KDG) (*18*, *19*). This catalytic promiscuity was predicted because of the presence of conserved residues required by MR for catalysis of the racemization reaction: the sequence of GlucD reveals homologues of both the general basic Lys-X-Lys motif and general basic His-Asp dyad presented in the active site of MR (*18*, *19*). We also were not surprised to discover that GlucD catalyzes the epimerization of D-glucarate and L-idarate (a 1,1-proton-transfer reaction analogous to that catalyzed by MR) in competition with the dehydration reactions.

The structure of the GlucD from *E. coli* was solved in the presence of Mg^{2+} and several active site ligands (*15*): 5-keto-4-deoxy-D-glucarate (KDG), the reaction product; 4-deoxy-D-glucarate (4DG), a substrate analogue that lacks the OH leaving group; and xylarohydroxamate, an analogue of the enolate anion intermediate. The structures of the KDG and 4DG ternary complexes with Mg^{2+} reveal an unexpected

[1] Abbreviations: 4Fgluc, 4-deoxy-4-fluoro-D-glucarate; DG, 4-deoxy-D-glucarate; GalD, D-galactonate dehydratase; GlucD, (*D*)-glucarate dehydratase; KDG, 5-keto-4-deoxyglucarate; MePEG, methyl ether polyethylene glycol; MLE I, muconate lactonizing enzyme I.; MR, mandelate racemase; OSBS, *o*-succinylbenzoate synthase; SBC, The Structural Biology Center at the Advanced Photon Source, Argonne, IL; HEPPS, *N*-[hydroxyethyl]piperazine-*N*′-[2-hydroxypropanesulfonic acid].

differences from the structure of the ternary complex of MR with the competitive inhibitor S-atrolactate and Mg^{2+} (*16*). First, in the GlucD structures, both carboxylate oxygens of KDG and 4DG are coordinated to the Mg^{2+}; in the MR structure, the carboxylate group of the inhibitor is a monodentate ligand of the Mg^{2+}. Second, based on the structure of the MR complex, we predicted that (1) Lys 207 of the Lys 205-X-Lys 207 motif would be positioned to abstract the proton from C5 of D-glucarate, and (2) His 339 of the His 339-Asp 313 dyad would be positioned to abstract the proton from C5 of L-idarate. The differing coordination geometries of the active site ligands to the Mg^{2+} ions in these structures reversed the stereochemical roles of the general acid/base catalysts in GlucD.

Our previous mutagenic and structural studies of GlucD focused on investigating the importance and catalytic importance of several functional groups, Lys 207 and His 339, the putative general acid/base catalysts, as well as Tyr 150 and Asp 366, two residues that are appropriately positioned to interact with and stabilize the enediolate intermediate (*15*). In the present manuscript, we identify the general acid catalyst that facilitates the vinylogous elimination of the 4-OH group from the enediolate anion intermediate.

MATERIALS AND METHODS

Site-Directed Mutagenesis and Protein Production. The N341D and N341L mutations were generated using the strategy used in our earlier studies (*15*).

4-*Deoxy-4-fluoro-D-glucarate (4-FGluc).* The methyl pyranoside 4-deoxy-4-fluoro-D-glucose was synthesized according to the literature procedure (*20*). The fluorosugar was oxidized with HNO_3 in the presence of $NaNO_2$ and purified as previously described for the preparation of L-idarate (*15*). ^{1}H NMR δ (D_2O) ~4.4 (1H, ddd, J = 46, 5.5, 3.0, H4), 4.21 (1H, dd, J = 20, 3.4, H5), 4.11 (1H, ddd, J = 18.1, 5.9, 2.6, H3), 3.96 (1H, d, J = 3.8, 2H). ^{19}F NMR (D_2O) δ 8.96 (dt, J = 46, 20).

Assays for Formation of KDG. The formation of KDG was assayed with semicarbazide to detect the α-keto acid product (*15*). For routine assays, an aliquot containing GlucD was added to 1 mL of 10 mM acid sugar (D-glucarate, L-idarate, or 4-FGluc) in 50 mM Tris-HCl buffer, pH 7.5, containing 5 mM $MgCl_2$, at 22 °C. The amount of α-keto acid product formed at various times was quantified by detection of its semicarbazone at 250 nm. Determinations of k_{cat} and K_m were performed by varying the concentration of the acid sugar.

Crystallization. The mutant proteins were dialyzed against 10 mM Tris-HCl (pH 8.0 at 4 °C), 5 mM $MgCl_2$. N341D and N341L were concentrated to 13.3 mg/mL and 14.5 mg/mL, respectively, and frozen as small pellets in liquid nitrogen. The frozen protein was stored at −80 °C and thawed daily for crystallization experiments. Crystals were grown with conditions used previously for wild-type protein (*15*). Crystals were grown by micro-batch experiments by combining equal volumes of protein with precipitant containing 14% MePEG 5000 (Aldrich Chemical Co., Milwaukee WI), 50 mM $MgCl_2$, 5% 2-propanol, and HEPPS, pH 8.0 at 4 °C. In some experiments, the protein was incubated with 5 mM D-glucarate prior to mixing with precipitant. The drops were then immediately macro-seeded with crystals of the same protein that were grown spontaneously in hanging drops using 10−16% MePEG 5000, 75 mM $MgCl_2$, 3% 2-propanol, HEPPS, pH 8.0, at 4 °C as the precipitant. In all experiments, wild-type and mutant proteins crystallized in multiple forms within the same wells. These two crystal forms are triclinic with unit cell a = 71 Å, b = 84 Å, c = 100 Å, α = 104°, β = 94°, γ = 114°, and an orthorhombic crystal form. For the wild-type protein, triclinic crystals, which were thicker and diffracted to 2.0 Å, were used for all experiments. The orthorhombic plates for the wild-type protein were always too thin for data collection.

Table 1: Data Collection and Processing Statistics

	N341D	N341L
space group	$P1$	$C222_1$
cell		
a (Å)	70.9	73.8
b (Å)	83.9	204.2
c (Å)	98.2	136.5
α (deg)	104.1	90
β (deg)	93.7	90
γ (deg)	113.1	90
maximum resolution (Å)	2.0	2.75
no. of observations	505 535	100 153
no. of reflections	132 888	24 649
completeness (%)	97.1 (89.0)[a]	91.1 (93.5)[a]
avg I	3679 (342)	2113 (271)
avg σ	196 (194)	106 (93)
R_{merge} (%)	6.2 (51.7)	7.1 (39.4)
Wilson B-factor (Å^2)	29.4	68.7

[a] Numbers in parentheses represent highest resolution shell, 2.1−2.0 Å for N341D and 2.85−2.75 Å for N341L.

With the Asn^{341} mutants, the protein crystallized more readily in the triclinic form in the presence of 5 mM D-glucarate while the orthorhombic form grew more commonly in the absence of substrate. The orthorhombic plates of the Asn^{341} mutants also grew thicker than the wild-type protein and were of sufficient quality for use in diffraction experiments. As discussed in the Results, this does not represent a change in the protein structure and the cause of this difference in behavior was not further investigated. A third crystal form of the N341D mutant protein also grew in the same wells and diffracted to 2.0 Å.

Data Collection and Structure Determination of the Triclinic Form of N341D. Data were collected at the SBC 19ID or 19BM for all experiments. For data collection, the crystals were transferred to an initial cryoprotectant containing 16% MePEG 5000, 0.1 M $MgCl_2$, 1% ethylene glycol, 5% 2-propanol, 5 mM D-glucarate, 50 mM HEPPS, pH 8.0 at 4 °C. After approximately 15 s in this solution, the crystal was transferred to the same solution containing 12% and finally 24% ethylene glycol. The crystals were in the 24% ethylene glycol cryoprotectant for 2 min prior to being flash frozen in a stream of nitrogen gas at −160 °C. After initial analysis and indexing with an in-house HiStar detector on a Rigaku generator, crystals were transported frozen to the SBC beamline 19ID for data collection.

Data were collected on a kappa goniostat in two 180° scans at ω = 0° and 180°. The data were collected with 1° frames and exposure times of 3 s. The two scans were integrated and scaled with HKL2000 (*21*). The data collection statistics are shown in Table 1. The CCP4 package (*22*) was used for Wilson scaling with TRUNCATE and 5% of the data were used for R_{free} calculations with UNIQUE and FREERFLAG.

Table 2: Final Refinement Statistics

	N341D	N341L
no. of reflections	130 650	24 617
R-factor (%)	19.8	21.3
R-free (%)	23.0	27.4
bond lengths (rmsd, Å)	0.005	0.007
bond angles (rmsd, deg)	1.18	1.27
no. of protein atoms	13 644	6801
no. of solvent atoms	1136	96
avg *B*-factor, protein (Å^2)	35.0	54.0
avg *B*-factor, solvent (Å^2)	40.0	37.6
avg *B*-factor, heteroatoms (Å^2)	40.2	50.8
PDB accession number	1JDF	1JCT

The molecular model was refined with difference Fourier methods using the wild-type protein as a starting refinement model. All solvent atoms, metal ions, ligands, and the side chain of Asn 341 were removed from the structure prior to refinement with CNS (version 0.5) (*23*). The initial crystallographic and free *R*-factors were 32.2 and 32.0%, respectively. Through a single round of simulated annealing, energy minimization, and individual *B*-factor refinement, the *R*-factors dropped to 25.1 and 28.3%. The structure was refined with iterative cycles of refinement and manual model building done with TURBO-FRODO (*24*). It became clear that the active site of all four subunits contained the product 5-KDG and this was built into the crystallographic model. The final refinement statistics are included in Table 2.

Data Collection and Structure Determination of the Orthorhombic Crystals of N341D and N341L. In the absence of D-glucarate, the predominant crystal form for the mutant enzymes was the orthorhombic crystal form. These crystals were transferred to cryoprotectant containing D-glucarate as described above. As the crystals were not grown in the presence of D-glucarate, the proteins only saw D-glucarate for approximately 4 min prior to being flash-frozen.

The data were also collected at SBC using the 19BM beamline and processed with HKL2000. Unlike the triclinic form of the enzyme which contained a complete tetramer in the asymmetric unit, the Matthews Coefficient for the orthorhombic crystals suggested that two subunits would be present in the active site with the remainder of the tetramer made up from a crystallographic 2-fold rotation axis. Molecular replacement was performed on the data with CNS using two subunits of the wild-type protein as a search model. A good solution was found which, when visualized with a graphics program, showed that the complete tetramer was formed around the crystallographic 2-fold rotation axis that is parallel to the *B* axis.

The data for the N341D mutant were processed and refined as described above for the triclinic crystal. At a crystallographic *R*-factor of 23.7%, it became apparent that the active site of both subunits contained 5-KDG and refinement was stopped. Data obtained from the N341L crystal were processed and scaled with HKL2000 and the CCP4 package (Table 1). The molecular replacement model derived from the N341D data were refined against the data for the N341L mutant using CNS. Initial electron density maps suggested that the active site contained glucarate and refinement continued to final crystallographic and free *R*-factors of 21.3 and 27.4%.

RESULTS AND DISCUSSION

The mechanism of the GlucD-catalyzed reaction involves at least three partial reactions, each requiring a catalytic functional group (Scheme 1): (1) general base-catalyzed abstraction of the α-proton to generate the stabilized enediolate anion; (2) general acid-catalyzed vinylogous elimination of the 4-OH group from the enediolate anion intermediate to generate an enol intermediate; and (3) general acid-catalyzed stereospecific tautomerization of the enol intermediate to generate the KDG product. While the identities of the general base-catalyst involved in the first partial reaction are known (Lys 207 for L-idarate and His 339−Asp 313 for D-glucarate; summary in the next section), the general acid catalysts involved the second and third partial reactions are unknown. The studies reported in this manuscript provide new functional and structural data that restrict the identities of the latter catalysts.

Previous Kinetic Analysis of the His 339 and Lys 207 Mutants of GlucD. In our previous study, we investigated the active site functional groups that initiate the dehydration reactions by abstracting the C5 protons from D-glucarate and L-idarate (*15*). As expected, based on the multistep nature of the overall reaction, these experiments indicated that the catalytic mechanism of GlucD was more complex than each base simply abstracting a proton from either D-glucarate or L-idarate. Mutation of either base dramatically reduced catalytic activity with both substrates (Table 3); if these functioned independently, it would have been expected that substitutions for His 339 would have prevented dehydration of D-glucarate but not L-idarate and that substitutions for Lys 207 would have the converse phenotype.

New Kinetic Studies of Lys 207 and His 339. The mutants of His 339 (H339N, H339Q, and H339A) and Lys 207 (K207Q and K207R) were impaired in the dehydration of both D-glucarate and L-idarate (Table 3). One possible interpretation is that one functional group acts the catalytic base, removing the α-proton from C5, and the other functional group acts as the general acid to protonate the 4-OH leaving group. The roles of these residues would be reversed with the epimeric D-glucarate and L-idarate substrates. This general base/general acid-catalyzed mechanism is similar to that envisaged for the dehydration reaction catalyzed by *o*-succinylbenzoate synthase reaction in which a Lys on one face of the active site abstracts the α-proton and the Lys on the opposite face delivers a proton to the leaving OH group (*14*).

To investigate the identity of the general acid in the GlucD-catalyzed reaction, 4-FGluc was synthesized and used as a substrate for wild-type GlucD and various active site mutants. In 4-FGluc, the leaving group is F^- rather than HO^-, so the dehydrofluorination reaction should not require general acid catalysis for facile vinylogous elimination of the 4-leaving group from the enediolate anion intermediate. This strategy was used previously to support the crystallographic identification of the general acid catalyst required in the dehydration reaction catalyzed by GalD, which, like GlucD, is related to MR by the presence of an active site His-Asp dyad. In the studies of GalD (*13*), structural and mutagenesis studies allowed the conclusion that the His 285−Asp258 dyad (at the ends of the seventh/sixth β-strands) is the general base that abstracts the α-proton; the structural work disclosed that

Scheme 1

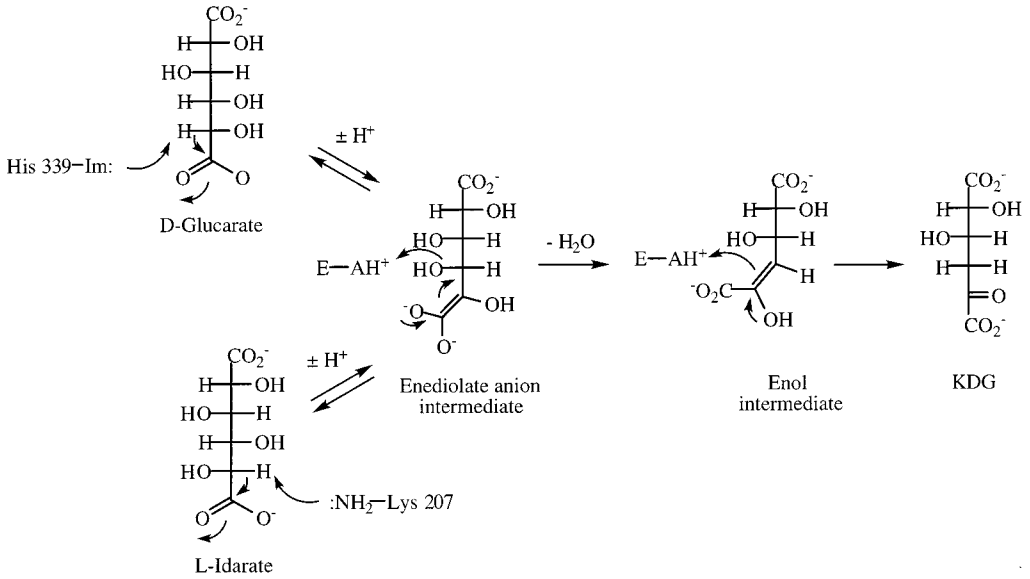


Table 3: Kinetic Analysis of Wild-Type and Mutant Enzymes[a]

	glucarate		idarate		4F-glucarate	
	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
WT[b]	35	580 000	35	210 000	44	13 000
H339Q[b]	0.020	422	0.020	139	0.030	30
K207R[b]	0.001	10	0.002	6.5	0.002	1.3
N341L	0.004	27	0.036	45	2.7	11 500
N341D	0.022	100	0.099	62	0.28	190

[a] The formation of the KDG product was quantitated as its semicarbazone as described in the text. All errors are less than 10%.[a] [b] The values presented for glucarate and idarate for these enzymes have been previously reported (*15*).

Table 4: Rates of Dehydration and Epimerization

	rate of dehydration (s^{-1})	rate of epimerization (s^{-1})	ratio
GlucD (D-glucarate)	35	23	0.66
GlucD (L-idarate)	35	29	0.83
N341L (D-glucarate)	4.4×10^{-3}	4×10^{-3}	1
N341L (L-idarate)	3.6×10^{-2}	2×10^{-3}	0.05
N341D (D-glucarate)	2.2×10^{-1}	2.2×10^{-2}	0.1
N341D (L-idarate)	9.9×10^{-3}	5.7×10^{-3}	0.58

[a] The rates of epimerization were quantitated by ^{1}H NMR spectroscopy as previously described (*19*). All errors are less than 10%.

His 185 was located on the opposite face of the active site (at the end of the third β-strand) and was appropriately positioned to facilitate the departure of the 4-OH group in an *anti*-elimination reaction. With D-galactonate as the substrate, the values for both k_{cat} and k_{cat}/K_m for the H185N and H185Q mutants were reduced to <0.01% of those of wild-type GalD. However, when 3-deoxy-3-fluoro-D-galactonate was used as substrate, the values of k_{cat} and k_{cat}/K_m for the H185N and H185Q mutants were ~10% those for the wild-type enzyme. This "rescue" by the 4-fluorinated substrate analogue was interpreted as evidence that His 185 is the general acid that facilitates the vinylogous elimination of the 4-OH leaving group in the GalD-catalyzed reaction.

Wild-type GlucD uses 4-FGluc as a substrate (Table 3) with the values of k_{cat} and k_{cat}/K_m comparable to those measured for D-glucarate and L-idarate. The K207R and H339Q mutants have dramatically reduced activities with 4-Fgluc (Table 3), with the reductions comparable to those observed for the "natural" diacid sugar substrates. This behavior illustrates that Lys 207 and His 339 play essential roles early in the reaction mechanism, prior to the elimination of the hydroxyl from the C4 position. The fact that 4-FGluc is not a good substrate for the K207R mutant suggests that Lys 207 is not the general acid catalyst for the departure of the 4-OH group from D-glucarate, as could be expected based on the both our earlier structural data as well as the data presented later in this manuscript. These kinetic results further suggest that both Lys 207 and His 339 are required to maintain the substrate in the proper orientation for proton abstraction by the general base on the opposite face of the active site.

Kinetic Analysis of the Asn 341 Mutants of GlucD. The only remaining residue that is positioned to participate in the departure of the 4-OH leaving group is Asn 341 (note that D-glucarate and L-idarate are epimers at C5, so the 4-OH leaving group is in the same position in both substrates). Asn 341 is located at the end of the seventh β-strand, on the C-terminal side of His 339. Two mutants, N341D and N341L, were constructed to study the role of Asn 341 in the dehydration reaction.

The reductions in both k_{cat} and k_{cat}/K_m for N341D and N341L with D-glucarate are comparable to those measured for the K207R and H339Q mutants (Table 3). The activities of the mutants were also quantified with 4-FGluc. While the N341D mutant is comparably inactive with D-glucarate, L-idarate, and 4F-Gluc, the N341L mutant retains nearly full activity with the substrate analogue. This observation suggests that while N341L is impaired in its ability to assist the departure of the 4-OH group, departure of the 4-F group proceeds unimpeded.

We also used ^{1}H NMR spectroscopy to measure the rates of epimerization of D-glucarate and L-idarate catalyzed by the N341D and N341L mutants (Table 4). Using either D-glucarate or L-idarate as substrate, wild-type GlucD catalyzes epimerization with a rate comparable to that measured for dehydration (*19*). If Asn 341 is involved in only departure

Table 5: RMS Deviations of the 441 Ordered Cα Carbons for a Representative Substrate of the Mutant and Wild-Type Protein Structures

	WT + 5KDG (Å)	N341L + glucarate (Å)
N341D + 5KDG	0.30	0.29
N341L + glucarate	0.30	

of the 4-OH group, the rates of epimerization catalyzed by the mutants might be expected to be similar to those measured for wild-type GlucD, despite the large reductions in k_{cat} for dehydration. Instead, we found that the rates of epimerization also were reduced and remained comparable to the rates of dehydration (Table 4). This observation is consistent with the functional group of Asn 341 influencing either the position or the acid−base properties of the spatially proximal His 339 rather than participating only in the dehydration reaction by facilitating the departure of the 4-OH group.

In this kinetic context, we determined high-resolution structures of N341D and N341L in the presence of active site ligands.

Crystal Structure of N341D in the Presence of 5KDG. The structures of the two Asn 341 mutant enzymes were determined to further investigate the role of Asn 341. The

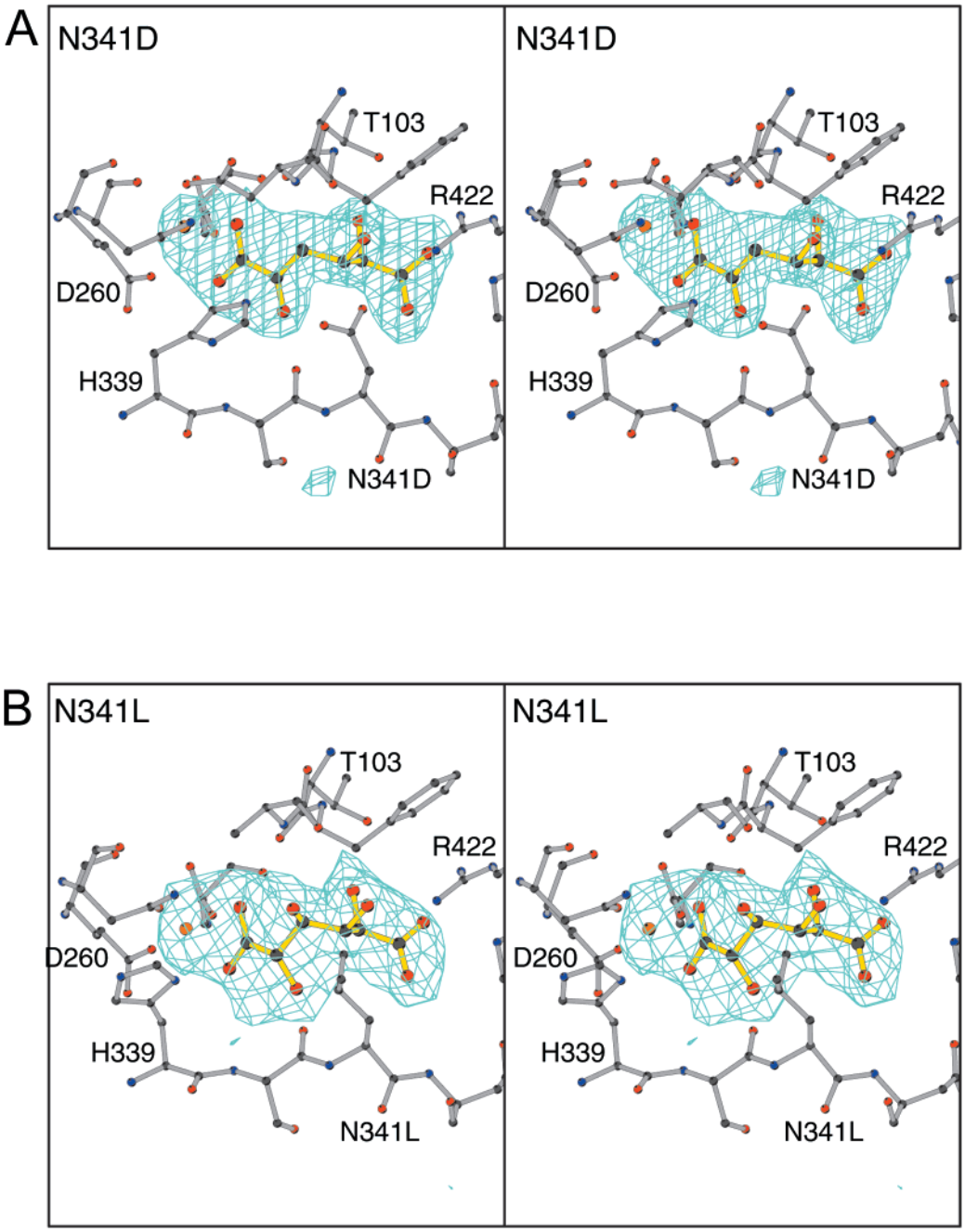


FIGURE 2: Active site density of mutant proteins. (A) The electron densities of the 5KDG and Mg^{2+} are shown for N341D mutant. The occupancies for the metal ion and product in all four subunits were set to 0.00 and subjected to a round of simulated annealing with CNS (*22*). The resulting $f_o - f_c$ map was contoured at 4σ. Also shown are several active site residues, His 339−Asp 341, Thr 103−Asp 105, His 32, Arg 422, and the three metal binding ligands, Asp 235, Glu 260, and Asn 289. B. The electron densities of the Mg^{2+} and D-glucarate molecule are shown for the N341L mutant. The map was generated in the same way and also contoured at 4σ. The same active site residues are included.

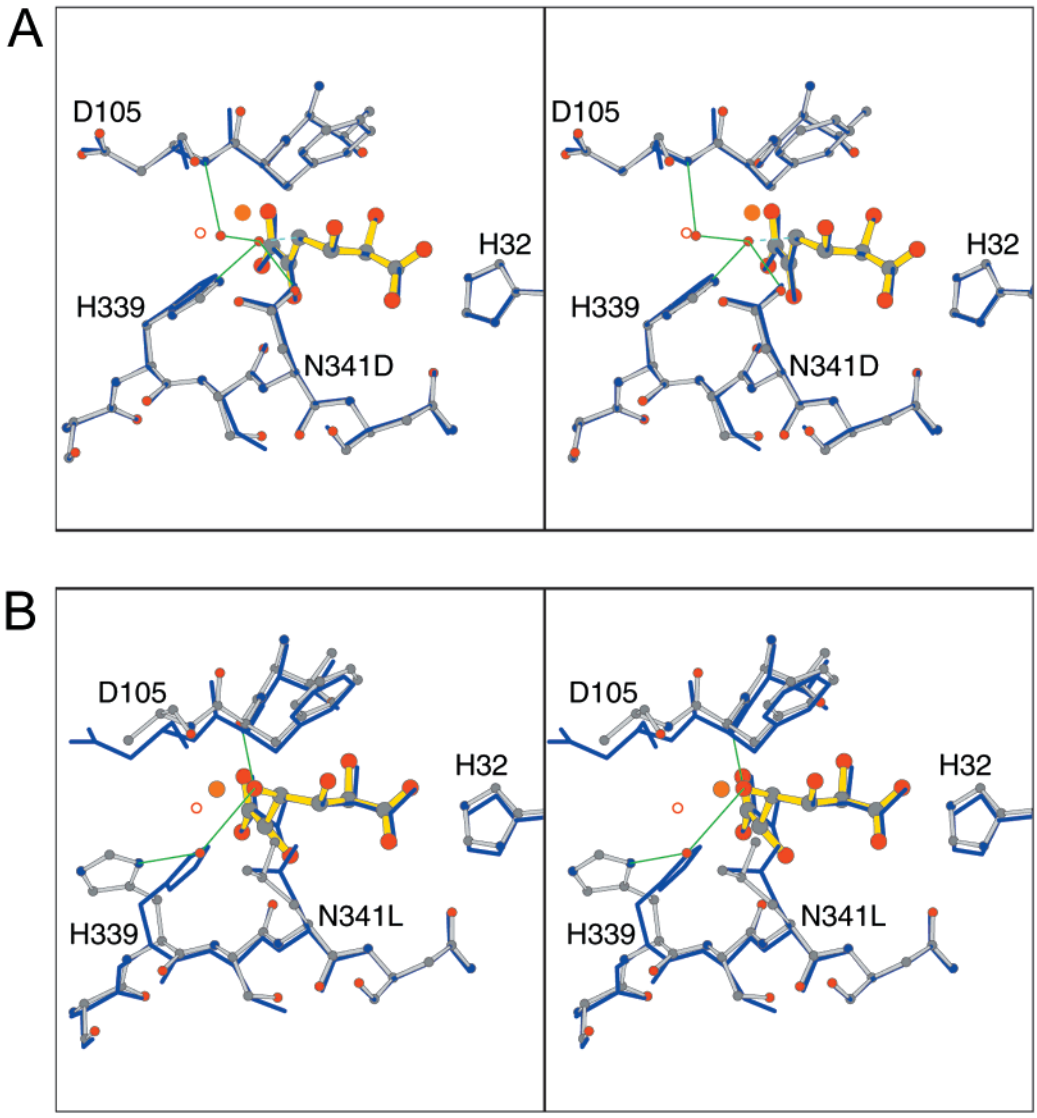


FIGURE 3: (A) Overlay of wild-type and N341D active sites. The active site of N341D (PDB accession number 1JDF) mutant protein is shown in ball-and-stick. The molecule of 5KDG is shown with yellow bonds. The wild-type protein is included for reference and shown with blue bonds (PDB accession number 1EC8). The water molecules and hydrogen bonding network observed in the N341D mutant are shown with green hydrogen bonds. The water molecule in the wild-type protein that hydrogen bonds with the main chain nitrogen of Asp 105 is shown as an open red sphere. The mutant protein structure contains an additional water molecule as is shown. No homologous water molecule is present in the wild-type structure. That the new water molecule is likely derived from the C4 hydroxyl is illustrated by the dashed blue line. (B) Overlay of wild-type and N341L active sites. The N341L mutant protein (PDB accession number 1JCT) is illustrated with ball-and-stick residues and the glucarate molecule is shown with yellow bonds. The wild-type protein again is shown with blue sticks representing the protein and active site contents. The rotation of the side chain of His 339 is apparent, as is the direct hydrogen bond between the C4 hydroxyl of D-glucarate and the carbonyl oxygen of Thr 103.

crystals of N341D were isomorphous with those of wild-type GlucD and were grown in the presence of 5 mM D-glucarate. Data were collected and processed as described in the Materials and Methods (Table 1) and refined to a final *R*-factor of 19.8%. The final model contained all four subunits which were complete except for the N-terminal four residues (Met 1, Ser 2, Ser 3, Gln 4). The following side chains are disordered: Gln 229 of chain A, Lys 60, Lys 158, Arg 169, Arg 190, Glu 214, Lys 386, and Glu 412 of chain B, Glu 214, and Lys 361 of chain C, and Thr 6, Se4 47, Glu 214, Lys 361, and Glu 412 of chain D. The only residue in the disallowed region of a Ramachandran plot is Thr 318 in all four subunits; the same geometry was observed in wild-type GlucD, although no structural reason is apparent to explain adoption of this strained conformation. The overall rmsd for the Cα atoms of the N341D structure is very low when compared to wild-type GlucD in the presence of 5KDG (Table 5).

Despite the low activity of the N341D mutant, the residual activity is sufficient to dehydrate D-glucarate when grown in the presence of this substrate (Figure 2A). The product was observed in the active sites of all four subunits. Comparison of the active sites of the wild-type and N341D showed very few differences that result from the substitution. The only apparent difference is the presence of a new water molecule in the active site of N341D that is positioned to form hydrogen bonds with the side chains of Asp 341 and His 339 and the carbonyl oxygen of Thr103 (Figure 3A). This water is likely to be derived from the leaving hydroxyl from C4 of the substrate (see below).

Because the N341D mutant cocrystallized with D-glucarate contained the product in the active site, crystals of N341D

also were grown in the absence of substrate. A second crystal form was obtained. Prior to freezing, the crystals were soaked in cryoprotectant containing 5 mM glucarate as described in the Materials and Methods. While the orthorhombic crystal form had been seen previously with the wild-type protein, the orthorhombic crystal form of N341D grew larger and more frequently than with the wild-type protein. On the basis of the unit cell, two subunits were located in the asymmetric unit, with the remaining 2-fold axis of the protein being provided by a crystallographic rotation axis. Two subunits of the wild-type GlucD structure were used as a search model for molecular replacement.

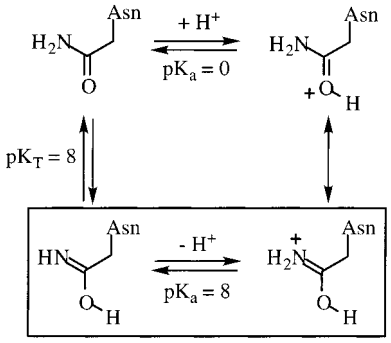
A solution was found which showed that the complete tetramer is made up around the 2-fold rotation parallel to the *B*-axis. This model was subjected to two cycles of crystallographic refinement and manual fitting. At a crystallographic *R*-factor of 23.7%, it became clear that the active site contained 5KDG, and refinement was discontinued. Like the triclinic form of the N341D mutant, a water molecule was hydrogen-bonded with the side chains of His 339, Asp 341, and the carbonyl oxygen of Thr 103. This water molecule was the only difference between the wild-type and N341D enzymes. This second crystal form of the N341D mutant serves to confirm the structure of the triclinic form which was described above. It also indicates that the structural change in the side chain of His 339 seen in the N341L mutant, described in the next section, is a result of the leucine mutation and not an artifact of the new orthorhombic crystal form.

The new crystal form also raises the possibility that the N341D mutant is chemically competent to turnover the substrate to product, since the product is observed in the active site after the brief soak in substrate. The reduction in activity that is observed (Table 3) may be due to a defect in product release. An impairment in product release would likely come from the inability of the flexible loop between α2 and α3 to open. The formation of the hydrogen bonding network between the side chains of His 339 and Asp 341 and the main chain nitrogen of Asp 105 on this flexible loop may trap the loop in the closed conformation.

Crystal Structure of the N341L Mutant Enzyme. The N341L mutant was crystallized in the absence of D-glucarate and soaked briefly in this substrate prior to freezing and data collection. The structure was solved for the orthorhombic crystal form by difference Fourier methods using the same solution as was obtained for the orthorhombic crystal form of the N341D protein. The active sites of both subunits contained D-glucarate, the substrate (Figure 2B).

Comparison of the structures of the wild-type and N341L proteins reveals a rotation of the side chain of His 339 in the mutant (Figure 3B). Although the kinetic constants for dehydrofluorination of 4F-Gluc by N341L (Table 3) could be explained by Asn 341 functioning as the general acid for departure of the 4-OH group, another possibility supported by the reduced rates of the epimerization reactions catalyzed by the mutants (Table 4) is that Asn 341 is required for proper positioning of His 339 so that it can act as the general acid. When D-glucarate is the substrate, His 339 is the general basic catalyst that initiates the reaction by abstraction of the α-proton from C5. Thus, the participation of His 339 as the general acid would require that this residue be optimally positioned to catalyze both initial proton abstraction as well as departure of the leaving group, partial reactions that occur at different locations in the substrate molecule.

On the basis of the structures for the N341D mutant in the presence of KDG, the water derived from the OH group that is eliminated from C4 of substrate can interact with the side chains of both His 339 and Asn 341; however, the identity of the residue that donates the proton to the 4-OH group is uncertain. The ability of N341L to catalyze dehydrofluorination of 4-FGluc but not D-glucarate could suggest that Asn 341 is the general acid. While this function is unlikely, given the pK_a of the carboxamide NH_s group [~17 (*25*)], proton delivery to the carboxamide carbonyl group will increase the acidity of the NH_s group. The following thermodynamic cycle (*26*) allows the conclusion that the pK_a of the NH_2 group could be depressed to ~8 assuming complete preequilibrium protonation [K_T is the equilibrium constant that relates the amide and iminol tautomers (*27*)]:



where K_T is the equilibrium constant that relates the amide and iminol tautomers; (*27*). With fractional protonation of the carboxamide carbonyl group (*25*), the apparent pK_a of the NH_2 group would be increased. Given the basicity of HO^- (the pK_a of water, its conjugate acid, is 15.7), a thermodynamically favorable protonation of the 4-OH group would be possible with partial protonation of the carboxamide carbonyl group. We note that the carboxamide carbonyl group of Asn 341 participates in a hydrogen-bonded network that includes several water molecules as well as the guanidinium group of Arg 38. Perhaps this network allows the carboxamide group of Asn 341 to function as a proton shuttle. Even if Asn 341 does not participate as a general acid in this reaction, the nucleophilicity of Asn residues in N-glycosylation reactions (*28*) might be explained by the local environment of the carboxamide group altering the pK_a of the NH_2 group by an analogous mechanism.

While this scenario cannot be discarded entirely, we believe that a more reasonable integrated interpretation of our kinetic and structural data is that Asn 341 is required to position His 339 so that it can function both as the general base for abstraction of the proton from C5 to generate the enediolate intermediate and, also, as the general acid in the vinylogous elimination of the 4-OH leaving group. In the dehydration of D-glucarate, His 339 would be protonated as a result of abstraction of the proton from C5; in the dehydration of L-idarate, His 339 would be protonated when the substrate is bound, thereby allowing the intermediate to partition between either vinylogous elimination of the 4-OH group and the observed epimerization.

The question that remains is how the N341L mutant is able to catalyze the dehydrofluorination of 4-FGluc with the His 339 residue in a position inappropriate for abstraction of the proton from the C5 carbon. Recall that the structure was determined for the N341L mutant with D-glucarate, not the fluorinated analogue, 4-FGluc. In the crystal structure, the position of the His 339 residue is certainly influenced by the presence of the hydrophobic leucine residue in the mutant position. With 4-FGluc as substrate, the bulky fluorine atom at C4 probably alters the position of the Leu 341 side chain. In this new position, the leucine may allow the His 339 to adopt the chemically competent conformation. In the 341D mutant, the diminished activity likely results from either an altered orientation of His 339, unfavorable electrostatic interactions between the carboxylate group and the departing 4-OH and 4-F groups, or inhibition of product release.

His 339 also is appropriately located to function as the general acid in the third partial reaction, stereospecific tautomerization of the enol intermediate to generate the KDG product. Our stereochemical studies of the GlucD-catalyzed reaction demonstrated that the 4-OH group is replaced with solvent hydrogen with retention of configuration. Given the structural data that reveal the proximity of only Asn 341 and His 339 to the 4-OH leaving group, the identification of the general acid for this partial reaction is subject to the same constraints on reactivity we have discussed for elimination of the 4-OH group: we also favor the participation of His 339 as the general acid catalyst for this partial reaction.

CONCLUSIONS

Both the kinetics and high resolution structural data for wild-type GlucD and its N341D and N341L mutants allow the reasonable conclusion that a single general acid/general base catalyst, His 339, is directly involved in all three partial reactions that result in *syn*-dehydration of D-glucarate: (1) general base-catalyzed abstraction of the α-proton from C5 to generate the stabilized enediolate anion; (2) general acid-catalyzed vinylogous elimination of the 4-OH group from the enediolate anion intermediate to generate an enol intermediate; and (3) general acid-catalyzed stereospecific tautomerization of the enol intermediate to generate the KDG product. While the *anti*-dehydration of L-idarate would involve the participation of Lys 207 as the general basic catalyst in the first partial reaction, His 339 would function as the general acid catalyst in the second and third partial reactions.

Thus, in contrast to the *anti*-dehydration reaction catalyzed by GalD that utilizes an additional general acid catalyst that is not shared with many other members of the enolase superfamily (*13*), the dehydration reactions catalyzed by GlucD involve only the canonical general acid/general base catalysts present in the active site of MR. This economy of functional groups, most "extreme" in the *syn*-dehydration of D-glucarate in which His 339 is used as the catalyst in all three partial reactions, emphasizes the functional plasticity of the $(\beta/\alpha)_7\beta$-barrel domain found in members of the enolase superfamily.

REFERENCES

1. 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) *Biochemistry 35*, 16489−501.
2. Babbitt, P. C., and Gerlt, J. A. (1997) *J. Biol. Chem. 272*, 30591−4.
3. Gerlt, J. A., and Babbitt, P. C. (1998) *Curr. Opin. Chem. Biol. 2*, 607−12.
4. Gerlt, J. A., and Babbitt, P. C. (2000) *Genome Biology 1* (5), reviews 0005.1−00005.10.
5. Gerlt, J. A., and Babbitt, P. C. (2001) *Ann. Rev. Biochem. 70*, 209−246.
6. Xiang, H., Luo, L., Taylor, K. L., and Dunaway-Mariano, D. (1999) *Biochemistry 38*, 7638−52.
7. Armstrong, R. N. (2000) *Biochemistry 39*, 13625−32.
8. Neidhart, D. J., Howell, P. L., Petsko, G. A., Powers, V. M., Li, R. S., Kenyon, G. L., and Gerlt, J. A. (1991) *Biochemistry 30*, 9264−73.
9. Wedekind, J. E., Reed, G. H., and Rayment, I. (1995) *Biochemistry 34*, 4325−30.
10. Helin, S., Kahn, P. C., Guha, B. L., Mallows, D. G., and Goldman, A. (1995) *J. Mol. Biol. 254*, 918−41.
11. Hasson, M. S., Schlichting, I., Moulai, J., Taylor, K., Barrett, W., Kenyon, G. L., Babbitt, P. C., Gerlt, J. A., Petsko, G. A., and Ringe, D. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 10396−401.
12. Gulick, A. M., Palmer, D. R., Babbitt, P. C., Gerlt, J. A., and Rayment, I. (1998) *Biochemistry 37*, 14358−68.
13. Wieczorek, S. W., Kalivoda, K. A., Clifton, J. G., Ringe, D., Petsko, G. A., and Gerlt, J. A. (1999) *J. Am. Chem. Soc. 121*, 4540−4541.
14. Thompson, T. B., Garrett, J. B., Taylor, E. A., Meganathan, R., Gerlt, J. A., and Rayment, I. (2000) *Biochemistry 39*, 10662−76.
15. Gulick, A. M., Hubbard, B. K., Gerlt, J. A., and Rayment, I. (2000) *Biochemistry 39*, 4590−602.
16. Landro, J. A., Gerlt, J. A., Kozarich, J. W., Koo, C. W., Shah, V. J., Kenyon, G. L., Neidhart, D. J., Fujita, S., and Petsko, G. A. (1994) *Biochemistry 33*, 635−43.
17. Larsen, T. M., Wedekind, J. E., Rayment, I., and Reed, G. H. (1996) *Biochemistry 35*, 4349−58.
18. Palmer, D. R., and Gerlt, J. A. (1996) *J. Am. Chem. Soc. 118*, 10323−10324.
19. Palmer, D. R., Hubbard, B. K., and Gerlt, J. A. (1998) *Biochemistry 37*, 14350−7.
20. Withers, S. G., MacLennan, D. J., and Street, I. P. (1986) *Carbohydr. Res. 154*, 127−44.
21. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*.
22. CCP4. (1994) *Acta Crystallogr., Sect. D 50*, 760−763.
23. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr., Sect. D 54*, 905−21.
24. Roussel, A., and Cabillau, C. (1991) *Silicon Graphics Geometry Partners Directory*, Silicon Graphics, Mountain View, CA.
25. March, J. (1992) *Advanced Organic Chemistry. Reactions, Mechanisms, and Structure*, John Wiley & Sons, New York.
26. Gerlt, J. A., Kozarich, J. W., Kenyon, G. L., and Gassman, P. G. (1991) *J. Am. Chem. Soc. 113*, 9667−69.
27. Schlegel, H. B., Gund, P., and Fluder, E. M. (1982) *J. Am. Chem. Soc. 104*, 5347−51.
28. Imperiali, B. (1997) *Acc. Chem. Res. 30*, 452−59.

BI010733B