The Structure of L-Ribulose-5-Phosphate 4-Epimerase: An Aldolase-like Platform for Epimerization[†,‡]

Yu Luo,[§] Jomy Samuel,[||] Steven C. Mosimann,[§] Jeffrey E. Lee,[§] Martin E. Tanner,*[,||] and Natalie C. J. Strynadka*[,§]

Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3, and Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1

Received June 18, 2001

ABSTRACT: The structure of L-ribulose-5-phosphate 4-epimerase from *E. coli* has been solved to 2.4 Å resolution using X-ray diffraction data. The structure is homo-tetrameric and displays C_4 symmetry. Each subunit has a single domain comprised of a central β-sheet flanked on either side by layers of α-helices. The active site is identified by the position of the catalytic zinc residue and is located at the interface between two adjacent subunits. A remarkable feature of the structure is that it shows a very close resemblance to that of L-fuculose-1-phosphate aldolase. This is consistent with the notion that both enzymes belong to a superfamily of epimerases/aldolases that catalyze carbon−carbon bond cleavage reactions via a metal-stabilized enolate intermediate. Detailed inspection of the epimerase structure, however, indicates that despite the close overall structural similarity to class II aldolases, the enzyme has evolved distinct active site features that promote its particular chemistry.

L-Ribulose-5-phosphate 4-epimerase (AraD, EC 5.1.3.4) catalyzes the interconversion of L-ribulose 5-phosphate (L-Ru5P)[1] and D-xylulose 5-phosphate (D-Xu5P) (Figure 1). It is one of three enzymes in the AraBAD operon that allow bacteria to utilize arabinose as an energy source by converting it into an intermediate in the pentose phosphate pathway (D-Xu5P) (*1*). Ru5P 4-epimerase is a homotetramer composed of four identical 25.5 kDa subunits and requires a divalent cation for activity (*2*−*6*).

Most racemases and epimerases act at stereocenters bearing a relatively acidic C−H bond and ultimately operate via a nonstereospecific deprotonation/reprotonation event directly at the site of inversion (*7*). The mechanism employed by Ru5P 4-epimerase is of interest since the substrate C-4 hydrogen is not adjacent to any carbonyl functionality and therefore is very nonacidic. For this reason, it is safe to assume that a simple proton-transfer mechanism is not at play with this enzyme. In addition, Ru5P 4-epimerase does not utilize NAD^+ as a cofactor (*2*), ruling out a nonstereospecific oxidation/reduction mechanism analogous to

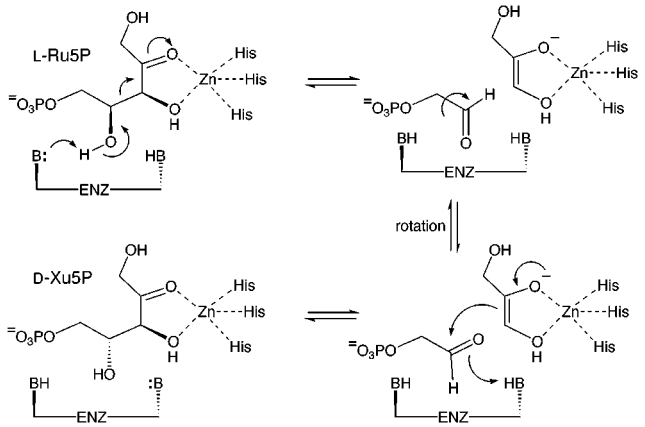

FIGURE 1: Mechanism of the reaction catalyzed by L-ribulose-5-phosphate 4-epimerase.

UDP-galactose 4-epimerase (*8*−*12*). Instead it appears that this enzyme employs a retroaldol/aldol mechanism in which carbon−carbon bond cleavage and re-formation ultimately lead to epimerization (Figure 1).

Consistent with this proposal, early mechanistic studies demonstrated that the epimerization proceeds without the detectable incorporation of oxygen or hydrogen atoms from the solvent (*13*, *14*). In addition, no primary kinetic isotope effect was observed during the epimerization of either D-[4-^{3}H]Xu5P or L-[4-^{2}H]Ru5P (*13*, *15*). More recently the reported sequence identity (26%) between the Ru5P 4-epimerase and the class II L-fuculose-1-phosphate (L-Fuc1P) aldolase has provided indirect support for a possible retroaldol/aldol mechanism in the epimerization event (*16*, *17*). The class II aldolases are found in bacteria and also use a divalent metal ion during catalysis (*18*, *19*). L-Fuc1P aldolase is a zinc metalloenzyme that is active as a homotetramer and catalyzes the condensation of L-lactaldehyde and dihydroxy-

[†] This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) (operating grant to M.E.T.) and the Canadian Institute of Health Research and the Howard Hughes Medical Institute to N.C.J.S.

[‡] X-ray coordinates have been deposited in the Protein DataBank (PDB accession code 1JD1; RCSB ID code RCSB013652).

* To whom correspondence should be addressed. N.C.J.S.: phone (604) 822-0789, fax (604) 822-5227, e-mail natalie@byron.biochem.ubc.ca. M.E.T.: phone (604) 822-9453, fax (604) 822-2847, e-mail mtanner@chem.ubc.ca.

[§] Department of Biochemistry and Molecular Biology, University of British Columbia.

[||] Department of Chemistry, University of British Columbia.

[1] Abbreviations: L-Ru5P, L-ribulose 5-phosphate; D-Xu5P, D-xylulose 5-phosphate; NAD^+, nicotinamide adenine dinucleotide; UDP, uridine diphosphate; L-Fuc1P, L-fuculose 1-phosphate; DHAP, dihydroxyacetone phosphate; PGH, phosphoglycolohydroxamate; EPR, electron paramagnetic resonance.

10.1021/bi0112513 CCC: $20.00
Published on Web 11/10/2001

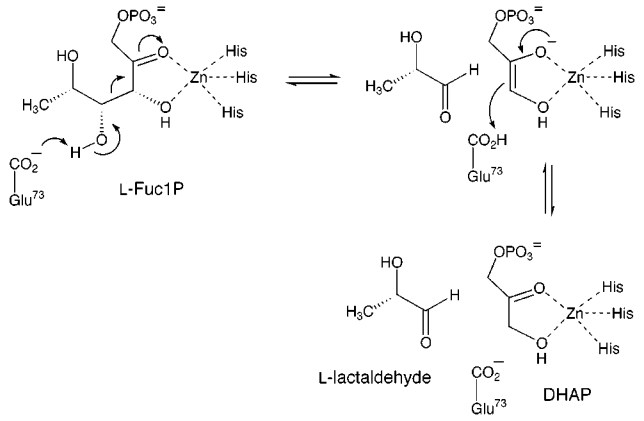

FIGURE 2: Mechanism of the reaction catalyzed by L-fuculose-1-phosphate aldolase.

acetone phosphate (DHAP). Structural information is available both on the free aldolase from *Escherichia coli* and on the complex with the enolate analogue phosphoglycolohydroxamate (PGH) (*16*, *20*, *21*), and recent mutagenesis studies have led to a mechanistic proposal for this enzyme (Figure 2) (*22*). The zinc in the free aldolase is ligated by a glutamate and three histidines (Glu73, His92, His94, and His155) arrayed in a roughly tetrahedral fashion. L-Fuc1P displaces the glutamate upon binding and coordinates to the zinc in a bidentate fashion via the carbonyl and the C-3 hydroxyl. The displaced Glu73 then serves to deprotonate the hydroxyl at C-4 and promotes carbon−carbon bond cleavage. This generates the metal-bound enolate of DHAP and L-lactaldehyde. The acidic form of Glu73 is then thought to protonate the enolate and yield the product DHAP.

The proposal that L-Fuc1P aldolase (and other class II aldolases such as L-rhamnulose 1-phosphate aldolase) and L-Ru5P 4-epimerase belong to a superfamily of aldolases/epimerases that utilize a similar retroaldol−aldol mechanism is further supported by recent mutagenesis studies involving conserved residues. Asp76, His95, His97, and His171 of the epimerase align with the corresponding metal-binding residues of the aldolase and presumably play a similar role. Indeed, the mutants D76N, H95N, and H97N showed a reduction in the value of k_{cat} and a reduced affinity for Zn^{2+} (*17*). In addition, it was shown that both the wild-type enzyme and D76N catalyzed the slow aldol condensation of glycolaldehyde phosphate and dihydroxyacetone to give a mixture of L-Ru5P and D-Xu5P (*17*). This was the first experimental evidence showing that the epimerase was capable of promoting carbon−carbon bond cleavage/formation and strongly suggests that the two chemical mechanisms are evolutionarily related. A recent report from Lee, Vu, and Cleland has provided definitive evidence in support of the aldolase-like mechanism (*23*). They observed primary ^{13}C isotope effects at both C-3 and C-4 during the epimerization of L-Ru5P. In addition, no primary deuterium isotope effects were detected at the same positions. These results are only consistent with a carbon−carbon bond cleavage mechanism and rule out other mechanisms involving deprotonation or dehydration events.

Additional work from the Cleland group pointed out some of the differences between the two enzymes (*24*). EPR studies with the Mn^{2+}-substituted enzyme in $H_2{}^{17}O$ showed that three waters were coordinated to the metal ion. This is notably different than with the aldolase in which no waters were ligated to the metal in the free enzyme. In addition, they found that mutagenesis of the C-terminal Tyr229 to Phe resulted in a 1000-fold drop in k_{cat}, implicating it as an important residue for catalysis. Mutagenesis studies with the aldolase, however, do not indicate that an analogous C-terminal residue is crucial for catalysis (*22*).

In this and the following paper (*25*), we explore the relationships between the two enzymes in greater detail. It is clear that the enzymes share a common mechanistic strategy in addition to certain structural features. Nevertheless, the differences in substrate structure, particularly the position of the phosphate, demand that key differences exist between the two. In this paper we present the structure of L-ribulose 5-phosphate 4-epimerase and provide a detailed comparison with the class II aldolases. In the accompanying paper, we employ site-directed mutagenesis to further probe the differences in their respective catalytic strategies (*25*).

MATERIALS AND METHODS

Purification and Crystallization. The L-Ru5P 4-epimerase from *E. coli* was overexpressed in *E. coli* and purified as described previously (*17*). The purified epimerase was crystallized using the hanging-drop vapor diffusion method. The crystals used for data collection were obtained using a precipitant similar to that previously reported (*26*). The drops were set up with an equal volume of ~10 mg/mL protein sample in the purification buffer and a precipitant composed of 4.0 M sodium formate and 0.1% *n*-octyl-*β*-D-glucoside. The square-shaped crystals grew to a typical size of 0.5 mm × 0.5 mm × 0.2 mm in a week.

Data Collection and Processing. The diffraction data were collected at the Brookhaven National Laboratory beamline X12C using a Brandeis-4 CCD detector system. The crystal was transferred to a cryoprotectant containing 4.0 M sodium formate and 25% glycerol, and was flash-cooled into a cryogenic stream operating at 100 K. The crystal diffracted to at least 2.2 Å. The data were processed to a resolution of 2.4 Å using HKL-1.96. 1 (*27*). The crystal belongs to the space group $P42_12$ as previously reported (*26*). It has slightly different unit cell dimensions: $a = b = 105.9$ Å and $c = 274.8$ Å. The composition of the cryoprotectant appeared to influence the cell parameters, particularly the *c* axis. The statistics of the data collection are listed in Table 1.

Structure Determination Using the Molecular Replacement Method. Processed data were imported to the CCP4 program package (*28*). Self-rotation function analysis revealed no convincing noncrystallographic symmetry (NCS). L-Ru5P 4-epimerase shows moderate sequence identity with the *E. coli* L-Fuc1P aldolase (26%), and both enzymes are tetrameric. A crystal packing density calculation showed that there are ~6 monomers per asymmetric unit. In addition, the fact that the cell dimension *c* is unusually large makes it reasonable to assume the epimerase tetramers pack along the crystallographic 4-fold axis. Starting from a trimmed aldolase tetramer model with its 4-fold axis aligned with the crystallographic *c*-axis, an AmoRe (*29*) translation search against 15−5 Å data was performed using an artificial one-dimensional rotation list with a 5° grid on the Eulerian angle α. The orientation of the aldolase monomer in the top solution was fitted to obtain a better tetramer search model.

Table 1: Data Collection and Refinement Statistics

Data Collection Statistics	
space group	$P42_12$
cell dimensions	$a = b = 105.9$ Å, $c = 274.8$ Å, $\alpha = \beta = \gamma = 90°$
res (Å)	2.40 (2.49−2.40)[a]
reflections	58293 (4187)
redundancy	3.64
complete	93.8 (68.8)
$I/\sigma I$	19.4 (4.1)
R-merge	0.056 (0.201)
Refinement Statistics	
model composition	6 × 223 = 1338 residues + 433 water molecules
resolution range	15.0−2.40 Å
reflections in working set	54929
reflections in test set	3283
R-factor	0.208
free R-factor for 10% data	0.241
rms deviation from ideality	
bond	0.0067 Å
angle	1.262°
Ramachandran plot	
non-glycine and non-proline residues in most favorable region	88.2%
outliers	none

[a] The numbers in parentheses are statistics for the highest-resolution bin.

Five more plausible solutions were found using the refined model. The final molecular replacement model gave a correlation coefficient of 0.498 and an R-factor of 0.46. The six tetramers pack along the crystallographic 4-fold with an alternate orientation and approximately equal spacing. In parallel with the molecular replacement trials, selenomethionine-incorporated protein was generated and crystallized isomorphously with the native protein. Multiwavelength anomalous diffraction data were collected at BNL, beamline X12C. A total of 46 sites corresponding to labeled methionines in the protein were identified by the software SOLVE (*30*), and subsequent maps were in complete agreement with our molecular replacement solution.

Phase Extension to 2.4 Å. Initial molecular replacement model phases and a 6-fold NCS were incorporated into the program DM (*31*). Phases were improved and gradually extended from 6 to 2.4 Å by 100 cycles of NCS-averaging, histogram-matching, and solvent-flattening. The DM-generated phases produced an easily traceable 2.4 Å electron density map. Despite the fact that the molecular replacement model did not include Zn^{2+}, the geometry around the Zn^{2+} is clearly recognizable in the map (Figure 3).

Model Building and Refinement. The electron density map was traced using the program XFIT (*32*). The model was refined with the program CNS version 0.9 (*33*) using 15−2.4 Å data. The NCS-restraints were initially applied to both main chain and side chain atoms, and were gradually released. With two additional cycles of remodeling peptide flips and side chain rotamers, the model was refined to an R-factor of 0.234 and a free R-factor of 0.261 using 6% flagged data. Water molecules were subsequently built into the model at peaks over 4.0σ in the F_o-F_c map. The tentative solvent molecules with refined B-factors lower than 60 Å^2 were retained. The final atomic model consisted of 223 × 6 = 1338 residues and 433 water molecules. It has an R-factor monitor of 0.208 and a free R-factor monitor of 0.241. No residues were found in disallowed regions of the Ramachandran plot as defined by PROCHECK (*34*). The zinc ions were refined with an occupancy of 1.0. The final B-factors of the zinc ions in each of the six monomers of the asymmetric unit are 31.3, 32.7, 29.0, 28.3, 41.1, and 69.2 Å^2, respectively.

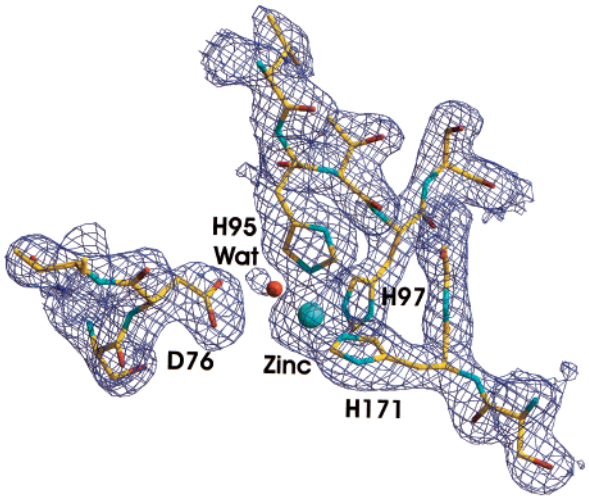

FIGURE 3: $2F_o-F_c$ electron density for the zinc-binding site of L-Ru5P 4-epimerase contoured at 1.5σ. The three conserved histidine residues ligate the zinc ion, while the adjacent Asp76 side chain carboxylate does not. The map was generated by XFIT (*32*) and rendered by Raster3D (*42*).

RESULTS

Tertiary and Quaternary Structure. In the structure of L-Ru5P 4-epimerase, the first 223 residues of each monomer are crystallographically ordered; however, the remaining 8 residues (224−231) are disordered and could not be located. The enzyme crystallizes as a homotetramer with C_4 symmetry reflecting the quaternary structure observed in solution studies (Figure 4). The solvent-accessible surface buried at the tetrameric interface is 2540 Å^2 per monomer, a value comparable to that usually observed with multimeric proteins. This accounts for 24% of the solvent-accessible surface of a subunit in the monomeric state.

Each monomer has a single domain of typical α/β fold (Figure 5A). A central β-sheet is formed from nine β-strands (b1−b9) and is predominantly antiparallel except between b7 and b8. The eight α-helices (a1−a8) of the structure form two layers on either side of the central β-sheet. The larger layer consists of five α-helices (a1, a3, a4, a7, a8), while the smaller layer consists of the remaining three helices (a2, a5, a6).

Comparison with the L-Fuc1P Aldolase Structure. Despite the moderate sequence identity (26%) between the epimerase and the aldolase, the two enzymes display strikingly similar tertiary and quaternary structures. As mentioned previously, each epimerase monomer has only 223 ordered residues, and the remaining 8 C-terminal residues are missing in its electron density map. Similarly, the aldolase (with or without a bound inhibitor) has 206 ordered residues and 9 disordered C-terminal residues (*16*, *20*, *21*). Both enzymes are tetrameric, and their individual monomers share a fold consisting of a central antiparallel β-sheet sandwiched by two layers of helices (Figure 5A,B). When the 2 monomeric structures were compared using the program ALIGN, 192 out of a

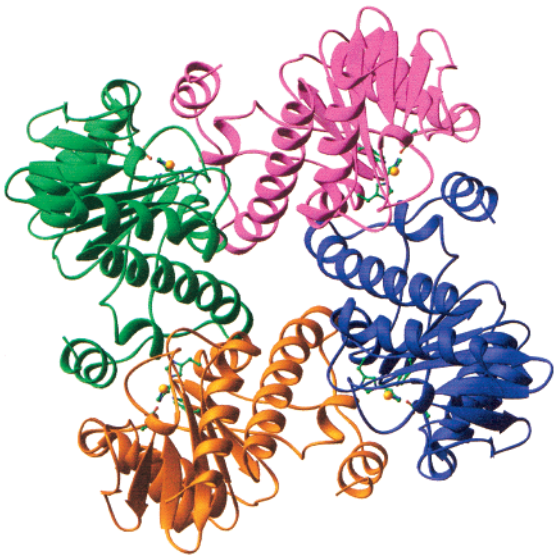

FIGURE 4: Quaternary structure of the L-Ru5P 4-epimerase homotetramer. The enzyme has C_4 symmetry. Each of the four monomers in the ribbon model are colored differentially. The four zinc cations are shown as gold spheres and are located at the interface of two adjacent monomers. The histidine ligands in each monomer are shown in ball-and-stick. The figure was generated with MOLSCRIPT (*43*) and rendered by Raster3D (*42*).

maximum of 206 (93%) Cα pairs aligned with a root-mean-square deviation of 1.52 Å. Among the aligned residues, 49 are identical in sequence, and 35 of these are found in the N-terminal half of the structures. Perhaps more importantly, when the tetrameric structures were compared, 192 × 4 = 768 (93%) Cα pairs aligned with a root-mean-square deviation of 1.89 Å. The comparable results between monomer and tetramer alignments indicate the two tetramers have strikingly similar quaternary architecture. This is readily apparent in a sequence alignment with the accompanying secondary structure displayed (Figure 6). All secondary structural elements in the smaller aldolase are conserved in the larger epimerase. The main difference between the two structures is that an additional short surface α-helix (a5, Asp134−Gln138) is observed in the small layer of helices (a2, a5, a6) in the epimerase. Other differences include the extension of the epimerase helix a6 (Tyr140−Ile158) by one turn as compared to the aldolase counterpart, and in the epimerase, α-helix a4 (Thr116−Asp120) replaces the topologically equivalent 3_{10}-helix of the aldolase (Tyr113−Ala118). This latter difference, although subtle, is important to note as helix a4 participates in formation of the active site and small changes in the disposition of this helix likely affect its contribution to catalysis in each of these enzymes (see below for further discussion).

Cation-Binding Site. The four active sites of the tetrameric epimerase are identified by the position of the catalytic zinc ions. As previous sequence analyses and mutagenesis studies have implied, the three conserved histidine residues (His95, His97, and His171) serve as ligands to the bound zinc ion (Figure 3) (*17*, *24*). His95 and His97 lie in a loop between b5 and a3, while His171 lies between b8 and b9. The observed distance between the Zn^{2+} and the three His NE2 atoms is 2.1, 2.1, and 2.0 Å for His95, His97, and His171, respectively. In the aldolase, the three histidines lie at topologically equivalent positions, and the corresponding

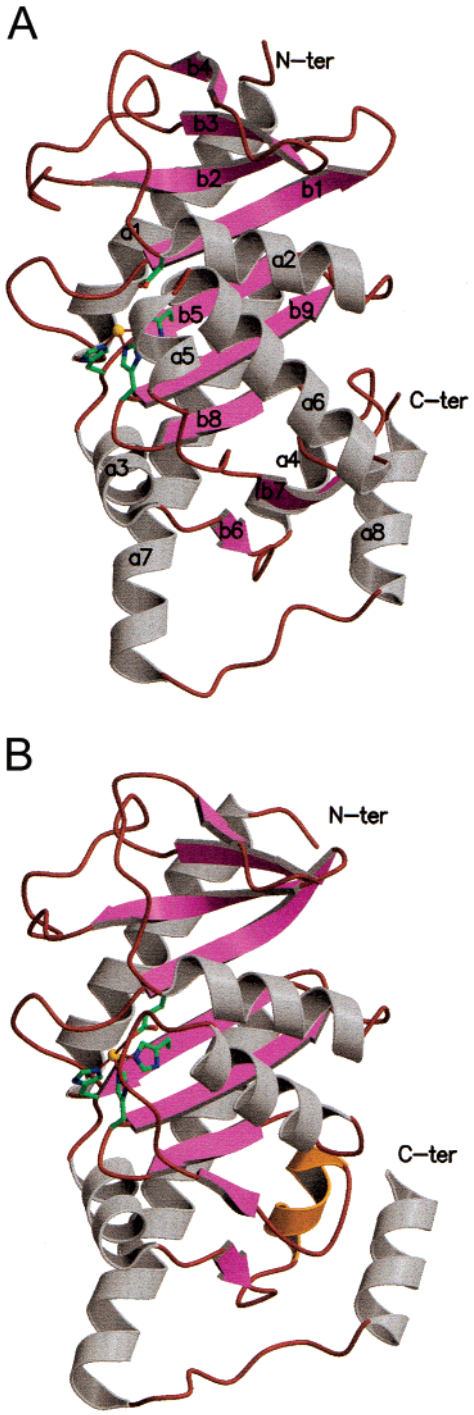

FIGURE 5: Ribbon diagrams of the monomer fold of (A) L-Ru5P 4-epimerase and (B) L-Fuc1P aldolase. Both enzymes have a central nine-stranded β-sheet flanked by one layer of α-helices at each side. The secondary structures of the epimerase are sequentially labeled and correspond to those shown in Figure 6. The β-strands, α-helices, and the 3_{10}-helix are colored in magenta, gray, and orange, respectively. The structures were superimposed and are viewed from the same direction. The zinc cation (gold sphere) is liganded by three conserved histidines in both structures, and additionally liganded by Glu73 in L-Fuc1P aldolase.

distances are 2.1, 2.1, and 2.0 Å for His92, His94, and His155, respectively. When comparing the two structures, all atoms of the three conserved His residues and the Zn^{2+} superimpose with a rms deviation of 0.45 Å. However, in the case of the epimerase, the three NE2−Zn^{2+}−NE2 angles

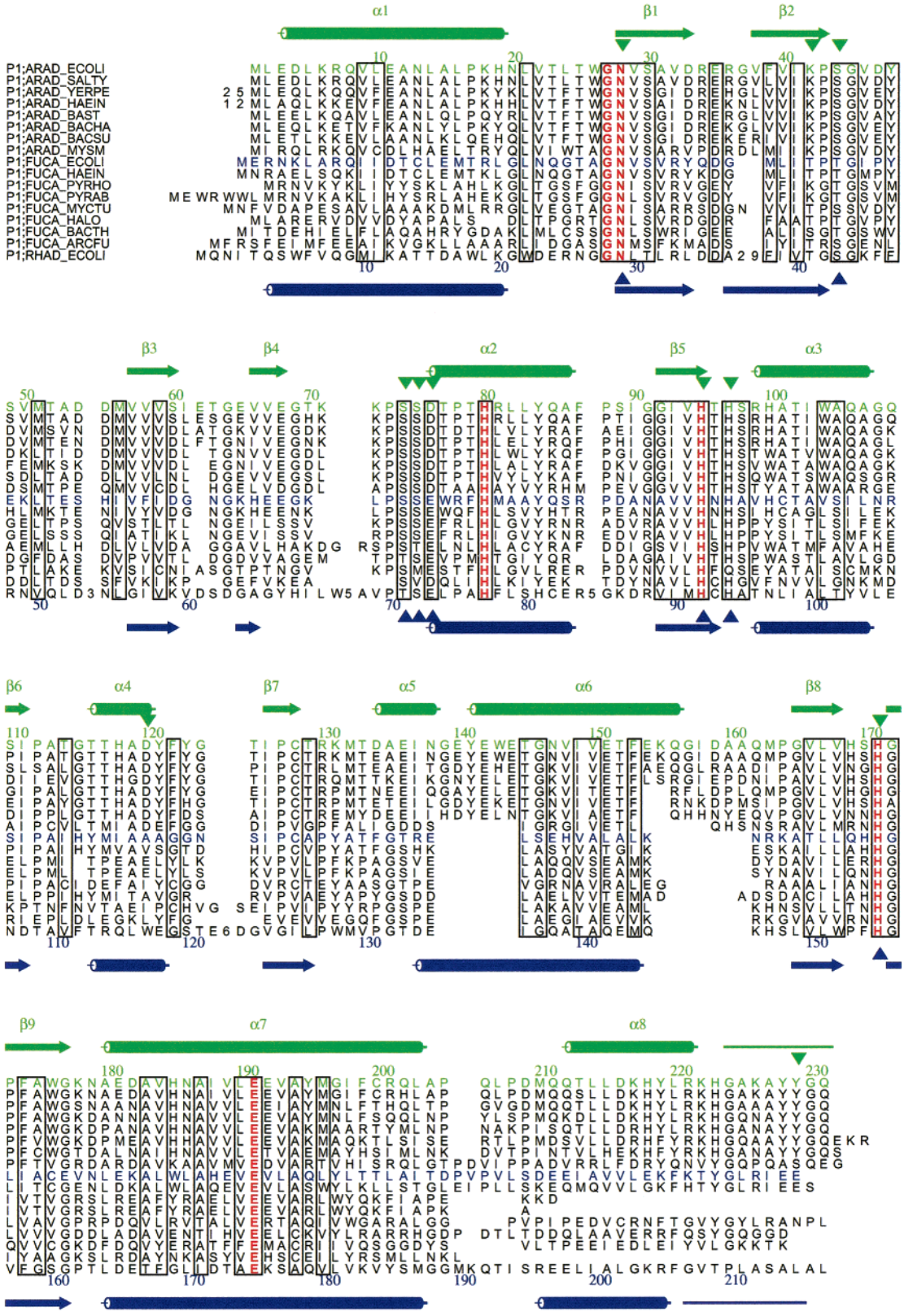

FIGURE 6: Multiple sequence alignment of L-Ru5P 4-epimerase and L-Fuc1P aldolase and their analogues. The alignment was obtained using Clustal W (*44*), and shown using Alscript (*45*). The *E. coli* epimerase sequence and its secondary structural elements are color-coded in green, the *E. coli* aldolase counterparts in blue. The residues involved in Zn^{2+} binding or phosphate binding, or which are possible general acid/base residues are marked with triangles. The six universally conserved residues are shown in red. The first eight sequences clearly clustered as epimerase homologues, while the last eight clustered as aldolase homologues.

are 100.7° (H95−Zn−H97), 96.5° (H95−Zn−H171), and 92.6° (H97−Zn−H171) as might be expected for a metal with octahedral geometry. This differs somewhat from the aldolase in which the corresponding angles are 101.6°, 106.5°, and 105.6°, respectively, and a tetrahedral arrangement of the ligands is adopted. Coordination numbers of 4

(tetrahedral), 5 (distorted trigonal bipyramidal), and 6 (distorted octahedral) have been previously observed in the structures of a wide number of protein/zinc complexes (*35*).

A striking difference between the epimerase and the aldolase is in regard to the nature of the remaining zinc ligand(s). In the aldolase, the fourth ligand is supplied by Glu73 which ligates the zinc in a bidentate fashion with distances of 1.9 and 2.4 Å to OE2 and OE1, respectively. In the epimerase, however, the topological equivalent, Asp76, lies 4.15 Å distant from the cation. The shortened side chain therefore prevents the carboxylate from acting as a ligand, and the remaining ligand(s) must be supplied by bound waters. In each of the epimerase monomers, an ordered water molecule is located midway between the zinc and G27 N (opposite to H171 NE2 by 172°; Figure 3). However, distances of this water to the zinc ion vary in each of the six monomers of the asymmetric unit (ranging from 2.5 to 3.8 Å), and though it is tempting to speculate that this represents a potential metal ligand, the coordinating distance is larger than typically observed for a zinc ion [2.0 Å, (*35*)]. No other liganding water molecules are directly observed in any of the six molecules of the asymmetric unit. It may be that the medium resolution of our structure is insufficient to provide clear electron density for additional liganding waters, and/or these waters are relatively disordered. Analysis of the structure indicates that considering steric and geometric constraints around the zinc ion there would be room for two additional water ligands, one coordinated between the conserved Asp76 side chain and the zinc, and one coordinated to the solvent-exposed face of the zinc, opposite to His95 NE2. The distance between Asp76 OD2 and the zinc is 4.7 Å, and the orientation of the carboxylate side chain is such that a water hydrogen-bonded to OD2 would be optimally positioned to act as a zinc ligand.

Substrate-Binding Site and Potential Catalytic Acid/Base Residues. The active site, as determined by the position of the essential zinc ion, is located at the interface of two adjacent monomers in the homotetramer. Amino acid side chains from both monomers protrude into the active site and could potentially play a role in catalysis. One monomer forms one side of the active site, which extends along the base of the central β-sheet, with contributions from the loop preceding helix a4 (specifically the conserved residues Pro73−Thr77), the loop following b2 (Pro43−Asp47), and the linker region between helix a5 and a6 (Gly139−Glu142). The adjacent monomer provides 2 helices (a4′ and a8′) to form the other side of the active site cavity (residues or secondary structural elements from the second monomer are denoted with a prime symbol). Notably, the α-helix a4′ and the C-terminal α-helix a8′ (and presumably several of the eight disordered C-terminal residues) extend into the active site of the adjacent monomer, providing several candidate amino acid side chains for catalysis (Figure 7). Collectively, a deep active site cleft is formed, which is relatively shielded from bulk solvent. Several conserved intermolecular interactions serve to stabilize the active site interface including a salt bridge from Glu142 to Arg221′, a hydrogen bond from the side chain hydroxyl of Tyr141 to His218′, a hydrogen bond from His97 ND1 and Thr117′ OG, and a number of intermolecular van der Waals and hydrophobic interactions.

Five residues in the epimerase (Asn28, Ser44, Gly45, Ser74, and Ser75) are highly conserved and are found at positions that are topologically equivalent to those of the known phosphate-binding residues of the aldolase (Asn29, Thr43, Gly44, Ser71, and Ser72; Figures 6 and 7). When comparing the two structures, all atoms of these five residues superimpose to an rms deviation of 0.53 Å. It is therefore reasonable to suspect that the seryl/threonyl β-hydroxyls, the asparaginyl γ-amide, and the glycyl α-amide form hydrogen bonds with the phosphate of the substrate in the case of the epimerase as well. The universally conserved Asn28 lies at the bottom of the pocket, and the remaining residues comprise the periphery. In addition, the epimerase positions the side chain of Lys42 at the bottom of the putative phosphate-binding pocket. This lysine is present in all of the epimerase sequences, but is absent in all of the aldolase sequences (including the *E. coli* aldolase) (Figure 6). It is largely buried in the active site, and the Lys42 NZ has a solvent-accessible surface of only ∼4 Å^2 as compared to a typical value of 40−50 Å^2 for an exposed lysine. The binding of a phosphate in this pocket would completely bury this residue, suggesting that a favorable salt bridge between the two could greatly assist substrate binding.

An important aspect of the aldolase catalysis is that substrate binding results in a substrate-induced fit that rearranges and immobilizes the loop containing residues 23−27 (Figure 7). This is clearly observed in a comparison of the structure of the free aldolase to that complexed with the inhibitor phosphoglycolohydroxamate (PGH). The equivalent loop in the epimerase, however, contains bulkier, hydrophobic residues (Leu24 and Trp26 as compared to Gly25 and Ala27 in the aldolase) which form a small hydrophobic core with hydrophobic side chains in the adjacent monomer (Ile188′, Val189′, and Tyr121′). It would be difficult to imagine that these residues could easily undergo an analagous reorganization as to that observed in aldolase without a major disruption of the active site. Indeed, a comparison of the epimerase structure and the inhibitor-bound aldolase structure shows that the main chain atoms of Thr23 to Asn28 superimpose on the corresponding aldolase atoms (Gln24 to Asn29) with an rms deviation of 0.52 Å. A similar comparison with the free aldolase structure, however, gives poorer agreement (rms deviation of 1.30 Å), indicating that the conformation of the free epimerase more closely resembles that of the bound aldolase, and no analagous structural rearrangement upon substrate binding would thus be required in the former enzyme. In addition, this segment of the native epimerase is relatively well-ordered since it has an average main chain *B*-factor of 31.0 Å^2, ∼4σ lower than the overall average (34.1 Å^2, rms 0.7 Å^2). The comparable segment in the native aldolase structure has an average main chain *B*-factor of 38.6 Å^2, ∼21σ higher than the overall average (17.4 Å^2, rms 1.0 Å^2).

Unlike the aldolase mechanism, the retroaldol/aldol mechanism for the epimerase presumably requires the presence of two residues that function as acid/base catalysts. One base deprotonates the C-4 hydroxyl of the substrate in the L-Ru5P to D-Xu5P direction, and the other deprotonates the C-4 hydroxyl in the D-Xu5P to L-Ru5P direction. It is conceivable that a single residue could accomplish both tasks; however, given the geometrical requirements necessary to reach both faces of the epimers, and the lack of precedence for "one-base" mechanisms in the racemase/epimerase literature (*7*, *36*), this seems unlikely. A survey of all the conserved side

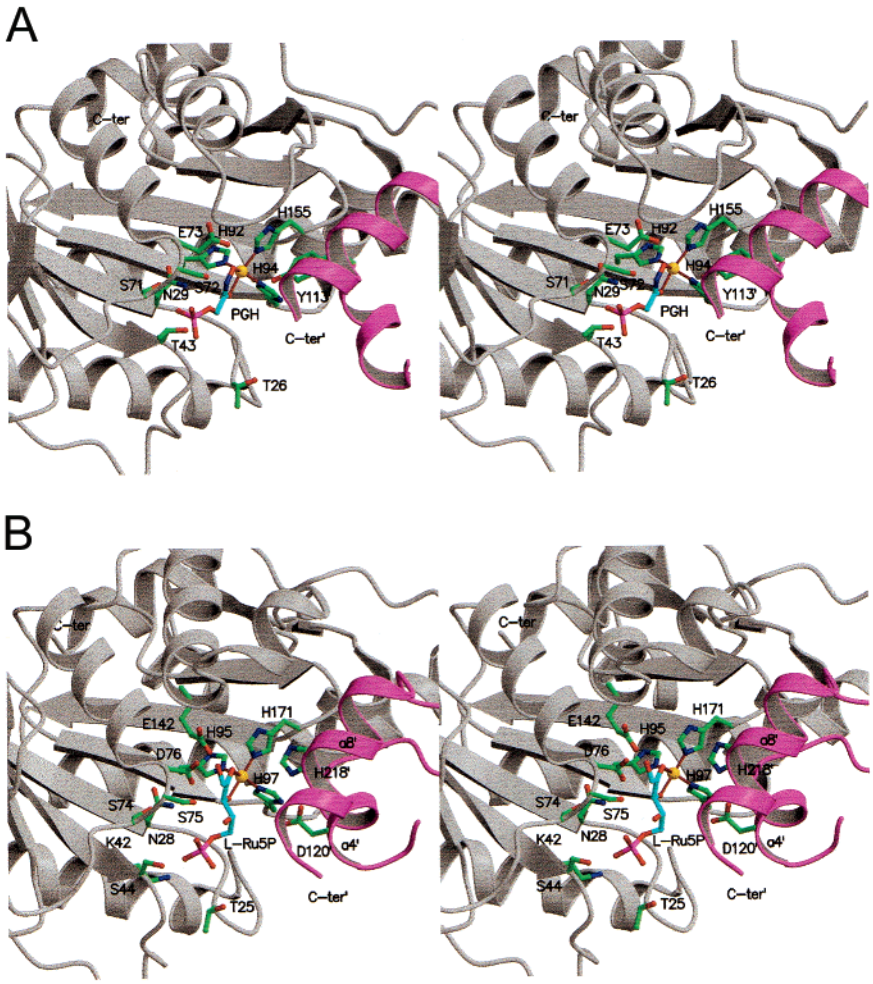

FIGURE 7: Stereo ribbon views of the substrate-binding sites of (A) L-Fuc1P aldolase with bound inhibitor PGH, based on the crystal structure, and (B) L-Ru5P 4-epimerase with bound substrate L-Ru5P, based on a molecular model of the complex (*25*). The secondary structural elements from two adjacent monomers that form the active site are shown in gray and magenta, respectively. Active site residues are shown in a ball-and-stick representation with green carbon, red oxygen, and blue nitrogen atoms. Residues that protrude from the second monomer are labeled with a prime symbol. The inhibitor and substrate are colored with cyan carbon, red oxygen, blue nitrogen, and magenta phosphate atoms. The inhibitor and substrate both provide two hydroxyls for coordination to the zinc ion. The 8 terminal residues that protrude from the C-terminal helix are disordered in the crystal structures of each enzyme but could obviously be positioned to extend across the active site.

chains in the vicinity of the epimerase active site reveals that Asp76 (analogous to the acid/base Glu73 of the aldolases) and Glu142 are reasonable candidates as well as Asp120′ and His218′ which are directed into the active site from α-helix a4′ and a8′ in the adjacent monomer of the homotetramer (Figure 7). It should also be pointed out that the last eight C-terminal residues (224−231) could not be located in the structure, yet the position of the preceding helix (212−222) and observation of a conserved glycine at position 223 (Figures 6 and 7) suggest this tail could be directed across the active site of an adjacent subunit. Thus, the conserved, titratable residues Tyr228′ and Tyr229′ within this region should also be considered candidates for catalysis.

DISCUSSION

In the past decade, it has become increasingly clear that the evolution of new catalysts is often guided by the specific chemical requirements of the reactions involved (*37*−*41*). Enzymes that catalyze seemingly different reactions and display low sequence homology may be evolutionarily related, and in many of these cases the relationships can be understood by considering the similarities in the mechanisms employed and the intermediates/transition states that must be stabilized. The close structural similarity between L-Ru5P 4-epimerase and L-Fuc1P aldolase clearly indicates that these enzymes have evolved from a common ancestor and belong to a superfamily of metal-dependent epimerases/aldolases. Both enzymes must catalyze the formation/cleavage of carbon−carbon bonds and stabilize an enolate ion through the use of a divalent cation.

While the two enzymes share many similarities, the subtle differences between them must account for the different catalytic activities. One such difference is seen in the residues employed as metal ligands. In the aldolase, three histidines and a glutamate serve as the ligands and bind the metal in a roughly tetrahedral fashion. The glutamate (Glu73) is displaced from the coordination sphere upon substrate binding and is then thought to serve as the key acid/base catalyst in the aldol reaction. In the case of the epimerase, Asp76 replaces the glutamate, and the shorter side chain precludes an interaction with the metal. An important consequence of this replacement is that the shorter residue is now unable to

serve as an acid catalyst and protonate the bound enolate intermediate (an event that would lead to unwanted aldolase activity).

Another consequence of the Asp substitution is that the remaining metal ligand(s) in the epimerase must be supplied by bound water. The exact number and position of the putative water ligand(s) are unclear in this structure; however, it is possible that three bound waters are allowed and an octahedral environment exists. This is consistent with the experimental work reported from the Cleland group in which the EPR spectrum of the Mn^{2+}-substituted enzyme was examined in $H_2{}^{17}O$ (*24*). The extent of the observed inhomogeneous broadening indicated that three water molecules were directly bound to the divalent cation. It should be mentioned that changing the metal from zinc to manganese could influence the coordination number; however, both Mn^{2+} and Zn^{2+} are able to accommodate 4 (tetrahedral) and 6 (octahedral) ligands in their coordination sphere (*35*).

The lack of bound substrate in the epimerase structure makes it difficult to definitively identify residues important for binding and catalysis. Nevertheless, the fact that the residues in the putative phosphate-binding pocket are strictly conserved between all aldolases and epimerases supports the notion that their function is similar (Figure 7). An additional lysine is present at the bottom of the epimerase phosphate-binding pocket that may increase the affinity toward the substrate. It is interesting to note that with the epimerase the value of K_M for L-Ru5P is 0.05 mM (*25*), whereas with the aldolase the value of K_M for L-Fuc1P is only 2.2 mM (*22*). Other structural features which may contribute to the increased apparent substrate affinity include the substitution of disordered water molecules for the bidentate zinc ligand, Glu73, of the aldolase (it is presumably more favorable for the two liganding oxygen atoms of the substrate to displace the disordered water molecules than the glutamate carboxylate, Figure 7). In addition, the active site of the epimerase is apparently "pre-organized" for substrate binding, and there is no requirement for the analogous substrate-induced rearrangement of residues 24−29 that occurs with the aldolase (*21*). Both these latter considerations would imply smaller energy costs, and lower K_m values for the binding of the substrate to the active site in L-Ru5P 4-epimerase and related enzymes.

A structural search for possible catalytic acid/base residues led to the conserved Asp76 and Glu142 (from the same subunit as the phosphate-binding site and metal ligands) as well as Asp120′ and His218′ (from helices a4′ and a8′ of the adjacent subunit, respectively; Figures 6 and 7B). Previous mutagenesis studies with the epimerase have ruled out Asp76 (analogous in sequence to the essential acid/base Glu73 of the aldolases) since a D76N mutant does not display the dramatic drop in the value of k_{cat} that would be expected (*17*). The residues Tyr228′ and Tyr229′ on the disordered C-terminal tail could also potentially play a catalytic role. One very likely candidate is Tyr229′ since the Y229F mutant has been reported to display only 0.1% of the activity present in the wild-type enzyme (*24*). It is important to note that the C-terminal tail in the FucA aldolase structure is also disordered (*16, 20, 21*). Mutagenesis studies with that enzyme suggest it may also be intimately associated with the active site, but imply a role in substrate binding rather than directly in catalysis (*22*). In addition, unlike the highly conserved C-terminal tail of the epimerases (Figure 6), the C-terminal residues of the aldolase family are highly variable in length and composition. Indeed, an intriguing aspect of the epimerase active sites relative to those of the aldolases is that although the packing of secondary structural units is relatively similar, the conservation of individual residues is quite different, particularly those which protrude from the adjacent monomer (α-helices a4′ and a8′ and the C-terminus). In the epimerases, these regions are highly conserved (Figure 6), underlining their potential importance in providing accurately positioned catalytic residues (Asp120′, His218′, Tyr229′, for example). On the other hand, the aldolases have no conservation in these same secondary structural elements, likely a reflection that these regions are not required for catalysis directly, although they may contribute to stabilization of the individual substrates in various species. In addition, subtle changes in geometry of these secondary structural elements are observed (for example, helix 4, which is a true α-helix in the epimerase and is a 3_{10}-helix in the aldolase) that influence the positioning of residues in the active site (for example, Asp120′ in the epimerase). Collectively these observations provide a striking example of adapting a common scaffold (in this case one involving an intermolecular interface) to the particular chemistry required.

The extremely close structural homology between the epimerase and the aldolase is consistent with the common mechanistic strategy employed by the enzymes. It is important to note, however, that the structures of the substrates differ significantly with respect to the position of the phosphate relative to the active center. Since phosphate recognition is likely to be a key factor in orienting the bound substrate, it is likely that these enzymes will utilize different catalytic residues during catalysis. These similarities and differences will be further investigated by mutagenesis studies in the accompanying paper (*25*).

ACKNOWLEDGMENT

We thank Robert Sweet for access to data collection at Beamline X12C at the NSLS, Brookhaven, NY.

REFERENCES

1. Lin, E. C. C. (1996) in *Escherichia coli and Salmonella* (Neidhardt, F. C., Curtis, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Resnikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., Eds.) pp 307−342, American Society for Microbiology, Washington, DC.
2. Deupree, J. D., and Wood, W. A. (1970) *J. Biol. Chem. 245*, 3988−3995.
3. Lee, N., Patrick, J. W., and Masson, M. (1968) *J. Biol. Chem. 243*, 4700−4705.
4. Mineno, J., Fukui, H., Yoshizumi, I., Ikunoshin, K., and Shinagawa, H. (1990) *Nucleic Acids Res. 18*, 6722.
5. Lee, N., Martin, G. R., Hamilton, E., and Fowler, A. (1986) *Gene 47*, 231−244.
6. Deupree, J. D., and Wood, W. A. (1972) *J. Biol. Chem. 247*, 3093−3097.
7. Tanner, M. E., and Kenyon, G. L. (1998) in *Comprehensive Biological Catalysis* (Sinnott, M., Ed.) Vol. II, pp 7−41, Academic Press, San Diego.
8. Tanner, M. E. (2001) *Curr. Org. Chem. 5*, 169−192.
9. Thoden, J. B., and Holden, H. M. (1998) *Biochemistry 37*, 11469−11477.
10. Frey, P. A. (1996) *FASEB J. 10*, 461−470.

11. Tanner, M. E. (1998) in *Comprehensive Biological Catalysis* (Sinnott, M., Ed.) Vol. III, pp 76−82, Academic Press, San Diego.
12. Frey, P. A. (1987) in *Pyridine Nucleotide Coenzymes* (Dolphin, D., Poulson, R., and Avramovic, O., Eds.) pp 461−477, Wiley, New York.
13. Davis, L., Lee, N., and Glaser, L. (1972) *J. Biol. Chem. 247*, 5862−5866.
14. McDonough, M. W., and Wood, W. A. (1961) *J. Biol. Chem. 236*, 1220−1224.
15. Salo, W. L., Fossitt, D. D., Bevill, R. D., Kirkwood, S., and Wood, W. A. (1972) *J. Biol. Chem. 247*, 3098−3100.
16. Dreyer, M. K., and Schulz, G. E. (1993) *J. Mol. Biol. 231*, 549−553.
17. Johnson, A. E., and Tanner, M. E. (1998) *Biochemistry 37*, 5746−5754.
18. Horecker, B. L., Tsolas, O., and Lai, C. Y. (1972) in *The Enzymes* (Boyer, P. D., Ed.) pp 213−258, Academic Press, New York.
19. Fessner, W.-D., Schneider, A., Held, H., Sinerius, G., Walter, C., Hixon, M., and Schloss, J. V. (1996) *Angew. Chem., Int. Ed. Engl. 35*, 2219−2221.
20. Dreyer, M. K., and Schulz, G. E. (1996) *Acta Crystallogr. D52*, 1082−1091.
21. Dreyer, M. K., and Schulz, G. E. (1996) *J. Mol. Biol. 259*, 458−466.
22. Joerger, A. C., Gosse, C., Fessner, W.-D., and Schulz, G. E. (2000) *Biochemistry 39*, 6033−6041.
23. Lee, L. V., Vu, M. V., and Cleland, W. W. (2000) *Biochemistry 39*, 4808−4820.
24. Lee, L. V., Poyner, R. R., Vu, M. V., and Cleland, W. W. (2000) *Biochemistry 39*, 4821−4830.
25. Samuel, J., Luo, Y., Morgan, P. M., Strynadka, N. C. J., and Tanner, M. E. (2001) *Biochemistry 40*, 14772−14780.
26. Andersson, A., Schneider, G., and Lindqvist, Y. (1995) *Protein Sci. 4*, 1648−1650.
27. Otwinowski, Z. (1993) in *Proceedings of the CCP4 study weekend: Data collection and processing* (Sawyer, L., Ed.) pp 29−33, SERC Daresbury Laboratory, Warrington, U.K.
28. Collaborative Computational Project, N. (1994) *Acta Crystallogr. D50*, 760−763.
29. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
30. Terwilliger, T. C., and Berendzen, J. (1997) *Acta Crystallogr. D53*, 571−579.
31. Cowtan, K. D. (1994) *Jt. CCP4 ESF-EACBM Newslett. Protein Crystallogr. 31*, 34−38.
32. McRee, D. E. (1999) *J. Struct. Biol. 125*, 156−165.
33. Brunger, A. T. (1998) *Acta Crystallogr. D54*, 905−921.
34. Laskowski, R. A., MacArthur, M. W., and Thornton, J. M. (1998) *Curr. Opin. Struct. Biol. 8*, 631−639.
35. Glusker, J. P. (1991) in *Advances in Protein Chemistry Metalloproteins: Structural Aspects*, pp 1−66, Academic Press, Inc., Harcourt Brace Jovanovich.
36. Sun, S., and Toney, M. D. (1999) *Biochemistry 38*, 4058−4065.
37. Babbitt, P. C., and Gerlt, J. A. (2001) in *Advances in Protein Chemistry* (Arnold, F. H., Ed.) pp 1−28, Academic Press, San Diego.
38. Gerlt, J. A., and Babbitt, P. C. (1998) *Curr. Opin. Chem. Biol. 2*, 607−612.
39. Babbitt, P. C., and Gerlt, J. A. (1997) *J. Biol. Chem. 272*, 30591−30594.
40. Babbitt, P. C., Mrachko, G. T., Hasson, M. S., Huisman, G. W., Kolter, R., Ringe, D., Petsko, G. A., Kenyon, G. L., and Gerlt, J. A. (1995) *Science 267*, 1159−1161.
41. Petsko, G. A., Kenyon, G. L., Gerlt, J. A., Ringe, D., and Kozarich, J. W. (1993) *Trends Biochem. Sci. 18*, 372−376.
42. Bacon, D. J., and Anderson, W. F. (1988) *J. Mol. Graph. 6*, 219−220.
43. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
44. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
45. Barton, G. J. (1993) *Protein Eng. 6*, 37−40.

BI0112513