Structure of 2-Amino-3,7-dideoxy-D-*threo*-hept-6-ulosonic Acid Synthase, a Catalyst in the Archaeal Pathway for the Biosynthesis of Aromatic Amino Acids[†,‡]

Mariya Morar,[§] Robert H. White,[||] and Steven E. Ealick*[,§]

Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, and Department of Biochemistry (0308), Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

Received May 15, 2007; Revised Manuscript Received July 13, 2007

ABSTRACT: Genes responsible for the generation of 3-dehydroquinate (DHQ), an early metabolite in the established shikimic pathway of aromatic amino acid biosynthesis, are absent in most euryarchaeotes. Alternative gene products, Mj0400 and Mj1249, have been identified in *Methanocaldococcus jannaschii* as the enzymes involved in the synthesis of DHQ. 2-Amino-3,7-dideoxy-D-*threo*-hept-6-ulosonic acid (ADH) synthase, the product of the Mj0400 gene, catalyzes a transaldol reaction between 6-deoxy-5-ketofructose 1-phosphate and L-aspartate semialdehyde to yield ADH. Dehydroquinate synthase II, the product of the Mj1249 gene, then catalyzes deamination and cyclization of ADH, resulting in DHQ, which is fed into the canonical pathway. Three crystal structures of ADH synthase were determined in this work: a complex with a substrate analogue, fructose 1,6-bisphosphate, a complex with dihydroxyacetone phosphate (DHAP), thought to be a product of fructose 1-phosphate cleavage, and a native structure containing copurified ligands, modeled as DHAP and glycerol. On the basis of the structural analysis and comparison of the enzyme with related aldolases, ADH synthase is classified as a new member of the class I aldolase superfamily. The description of the active site allows for the identification and characterization of possible catalytic residues, Lys184, which is responsible for formation of the Schiff base intermediate, and Asp33 and Tyr153, which are candidates for the general acid/base catalysis.

Bacteria and eukaryotic organisms utilize an established pathway, generally known as the shikimic acid pathway, for the biosynthesis of aromatic amino acids, the *p*-aminobenzoic acid moiety of folate, and various quinone precursors (*1*). This pathway begins with a condensation of erythrose 4-phosphate with phosphoenolpyruvate to yield 3-deoxy-D-*arabino*-2-heptulosonate 7-phosphate which is then cyclized to 3-dehydroquinate (DHQ)[1] (Scheme 1) (*2*, *3*). Euryarchaea and crenarchaea are the two major subdivisions of archaea, differentiated on the basis of DNA structure and transcription machinery (*4*, *5*). The euryarchaea do not contain the canonical genes leading to DHQ; however, the rest of the genes responsible for the shikimic acid pathway except one (*6*) are present (*7*−*9*).

Euryarchaea mainly include methanogens, halophiles, and sulfur-metabolizing thermophiles. Recent work on a euryarchaeotic methanogen, *Methanocaldococcus jannaschii*, showed that rather than utilizing nonorthologous genes for the enzymes responsible for DHQ production, the reactions leading to this intermediate are different (Scheme 1) (*10*). Here, the first step is a transaldolase reaction between 6-deoxy-5-ketofructose 1-phosphate (DKFP) and L-aspartate semialdehyde with an elimination of hydroxypyruvaldehyde phosphate to yield 2-amino-3,7-dideoxy-D-*threo*-hept-6-ulosonic acid (ADH) catalyzed by the Mj0400 gene product, ADH synthase (ADHS) (*11*). The second step is a NAD-dependent oxidative deamination of ADH followed by cyclization to yield DHQ, catalyzed by the Mj1249 gene product.

To further characterize the alternative steps of the biosynthetic pathway for aromatic amino acids in euryarchaea, we determined crystal structures of three *M. jannaschii* ADHS complexes. These structures allowed for the description of the overall architecture of the enzyme and that of the active site. Comparison of the ADHS with a closely related archaeal fructose 1,6-bisphosphate aldolase (FBPA) from *Thermoproteus tenax* (*12*) as well as the structures of other class I aldolases gives insight into the reaction mechanism catalyzed by this novel enzyme.

[†] This work was supported by NIH Grant RR15301 to S.E.E. and National Science Foundation Grant MCB-0231319 to R.H.W. S.E.E. is indebted to the Lucille P. Markey Charitable Trust and the W. M. Keck Foundation.

[‡] The coordinates of the ADH synthase structures have been deposited in the Protein Data Bank as entries 2QJG for the ADHS−F1,6P complex, 2QJH for the ADHS−DHAP complex, and 2QJI for the ADHS−DHAP−GLYC complex.

* To whom correspondence should be addressed: Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853. Telephone: (607) 255-7961. Fax: (607) 255-1227. E-mail: see3@cornell.edu.

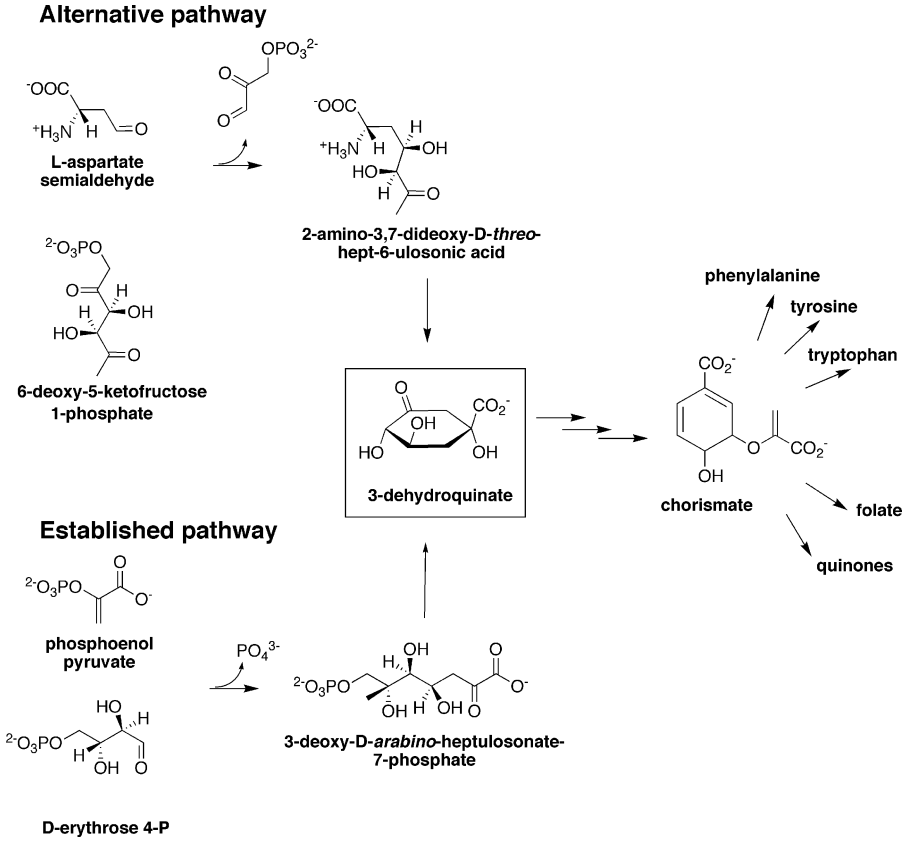
[§] Cornell University.

[||] Virginia Polytechnic Institute and State University.

[1] Abbreviations: ADH, 2-amino-3,7-dideoxy-D-*threo*-hept-6-ulosonic acid; ADHS, 2-amino-3,7-dideoxy-D-*threo*-hept-6-ulosonic acid synthase; Mj, *Methanocaldococcus jannaschii*; Tt, *Thermoproteus tenax*; FBPA, fructose 1,6-bisphosphate aldolase; DHQ, 3-dehydroquinic acid; DKFP, 6-deoxy-5-ketofructose 1-phosphate; ASA, L-aspartate semialdehyde; DHAP, dihydroxyacetone phosphate; F1,6P, fructose 1,6-bisphosphate; F1P, fructose 1-phosphate; PEP, phosphoenolpyruvate; GLYC, glycerol; CNS, Crystallography and NMR System; NCS, noncrystallographic symmetry; NBHA, *o*-(4-nitrobenzyl)hydroxylamine; CE, collision energy; RAMA, rabbit muscle aldolase; PwTIM, *Pyrococcus woessei* triosephosphate isomerase; StYiht, *Salmonella typhimurium* protein of unknown function.

10.1021/bi700934v CCC: $37.00
Published on Web 08/22/2007

Scheme 1



EXPERIMENTAL PROCEDURES

Cloning of mj0400. Cloning of *mj0400* to yield the pET-11 plasmid was previously described (*10*). The pET-28a construct of the ADHS was prepared by using standard blunt ligation subcloning procedures at the Protein Facility of the Department of Chemistry and Chemical Biology, Cornell University.

Expression and Purification of ADHS. The pET-28a plasmid was transformed into the *Escherichia coli* B834-(DE3) cell line and plated on a kanamycin resistant agar plate. Single colonies from the plate were used to innoculate 5 mL of LB medium starter culture supplied with 35 μg/mL kanamycin, and the starter cultures were grown overnight at 37 °C with shaking. The starter culture was then used to inoculate a 1 L culture of LB medium supplied with kanamycin and grown at 37 °C with shaking until an OD of 0.6 was reached. The cells were then induced with 100 μM β-D-thiogalactopyranoside for 2 h. The cells were harvested by centrifugation (8700*g* for 10 min) using a Beckman centrifuge. The cell pellet was frozen at −80 °C until purification.

The cell pellet was thawed in a lysis buffer containing 300 mM NaCl and 100 mM sodium phosphate (pH 7.5). The cells were then lysed via sonication using a sonic dismembrator (model 550) from Fisher Scientific, with a pulse of 5 s followed by cooling for 5 s over 6 min. The cell debris was then removed by centrifugation (35000*g* for 45 min). The cleared lysate was loaded onto a 2 mL Ni-NTA affinity agarose column from QIAGEN, equilibrated with the lysis buffer, at a flow rate of 1 mL/min. The column was washed with 100 mL of wash buffer containing the lysis buffer and 5 mM imidazole. The protein was then eluted in 3 mL fractions using the elution buffer, which was the lysis buffer containing 250 mM imidazole. The protein concentration during the elution was monitored using Coomassie Plus Protein Assay Reagent (Pierce).

The N-terminal polyhistidine tag was cleaved using biotinylated thrombin from Novagen, via dialysis of ADHS in the presence of thrombin using 10 kDa molecular mass cutoff Snakeskin Pleated dialysis tubing from Pierce overnight at 4 °C with the ratio of 100 μg of protein/unit of thrombin. The thrombin buffer consisted of 50 mM Tris (pH 8.2), 300 mM NaCl, 2 mM DTT, 2.5 mM $CaCl_2$, 2.5 mM $MgCl_2$, and 2% glycerol. The thrombin digestion was monitored by SDS gel analysis and proceeded to approximately 70% completion. The digestion was quenched by removing thrombin from the protein solution via a streptavidin−agarose capture column from Novagen. The ADHS solution was then buffer exchanged using an Econo-Pac 10DG desalting column (Bio-Rad) into 10 mM Tris (pH 7.5) and concentrated to 20 mg/mL using an Amicon Ultra centrifugal filter device with a 10 kDa molecular mass cutoff. This solution was used for crystallization experiments.

Crystallization of ADHS. The vapor diffusion hanging drop method was used; drops contained 2 μL of protein solution combined with 2 μL of well solution. Hampton and Wizard sparse matrix screens were used to obtain initial crystallization hits. Optimized crystallization conditions consisted of 4−8% 1,4-butanediol and 0.1 M acetate buffer (pH 4.2−4.3). The crystals grew in 1−2 days and were flash-frozen in liquid nitrogen with 20% glycerol and mother liquor serving as the cryoprotectant for data collection. Prior to freezing, a set of crystals was soaked for 1 h with either 10 mM fructose 1-phosphate (F1P) or fructose 1,6-bisphosphate (F1,6P) in mother liquor. ADHS crystallized in triclinic space

Table 1: Summary of Data Collection and Processing Statistics[a]

	ADHS−F1,6P	ADHS−DHAP	ADHS−DHAP−GLYC
resolution (Å)	2.6	2.6	2.8
space group	*P*1	*P*1	*P*1
wavelength (Å)	0.9792	0.9792	0.9792
a (Å)	94.45	94.49	94.14
b (Å)	102.77	101.87	103.54
c (Å)	156.41	154.16	153.99
α (deg)	89.40	90.31	90.05
β (deg)	85.83	86.70	87.97
γ (deg)	82.05	82.45	82.04
no. of reflections	316793	338417	263608
no. of unique reflections	169078	138057	132360
I/σ	12.0 (2.6)	11.5 (2.6)	6.3 (2.0)
R_{sym}[b]	6.1 (28.0)	7.2 (38.3)	10.5 (49.0)
redundancy	1.9 (1.7)	2.5 (2.4)	2.0 (1.9)
completeness (%)	95 (79)	90 (96)	93 (89)

[a] Values for the highest-resolution shell are given in parentheses. [b] $R_{sym} = \sum\sum_i|I_i - \langle I\rangle|/\sum\langle I\rangle$, where $\langle I\rangle$ is the mean intensity of the N reflections with intensities I_i and common indices h,k,l.

Table 2: Refinement Statistics

	ADHS−F1,6P	ADHS−DHAP	ADHS−DHAP−GLYC
resolution (Å)	50−2.6	50−2.6	50−2.8
no. of non-hydrogen atoms	41043	40024	41213
no. of protein atoms	40046	39432	40727
no. of ligand atoms	380	180	300
no. of water atoms	617	412	186
no. of reflections in refinement	159934	131371	125135
no. of reflections in test set	8021	13111	12530
R[a]	20.4	20.3	24.3
R_{free}[b]	24.4	24.4	29.6
rmsd for bonds (Å)	0.008	0.008	0.008
rmsd for angles (deg)	1.4	1.3	1.3
average *B* factor (Å^2)	30	53	30
Ramachandran plot			
most favored (%)	91.3	91.8	88.6
additionally allowed (%)	8.7	8.2	11.4
disallowed (%)	0	0	0

[a] R factor $= \sum_{hkl}||F_{obs}| - k|F_{cal}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{cal} are observed and calculated structure factors, respectively. [b] In R_{free}, the sum is extended over a subset of reflections (10%) that were excluded from all stages of refinement.

group *P*1 with 20 protomers per asymmetric unit and 50% solvent content.

Data Collection and Processing. All of the data reported here were collected at the 24-ID-C beamline of the Advanced Photon Source at Argonne National Laboratory (Argonne, IL) using a Quantum 315 detector with a crystal to detector distance of 400 mm, an oscillation step of 0.5°, and an exposure time of 1 s. The HKL2000 program suite was used to integrate and scale all data (*13*). The approximate triclinic unit cell dimensions for the three ADHS crystals were as follows: $a = 94.4$ Å, $b = 102.7$ Å, $c = 154.9$ Å, $\alpha = 90.1°$, $\beta = 86.8°$, and $\gamma = 82.2°$. Final data processing statistics are listed in Table 1.

Structure Determination and Refinement. The structure of ADHS was determined by molecular replacement using the structure of F1,6P aldolase from *T. tenax* (TtFBPA) as a search model [PDB entry 1OJX(*13*)]. The search model was 35% identical and 55% similar in sequence to MjADHS. The Crystallography and NMR System (CNS) program suite (*14*) was used to calculate the rotation and translation functions with a 4 Å resolution cutoff and one pentamer of TtFBPA used as a search model. One pentamer of TtFBPA was used to define the origin; the positions of the three remaining pentamers were located by performing three translational searches. To improve the quality of the electron density map, density modification was performed in CNS followed by 20-fold noncrystallographic symmetry (NCS) averaging using RAVE (*15*). The averaged density-modified map was then used to build the initial model. The subsequent model refinement was conducted in CNS with several rounds of rigid body refinement, annealing, and *B* factor refinement, followed by manual model building in COOT (*16*). NCS restraints were applied to the structures during the refinement in CNS. The averaged composite omit map was used when refining the models manually. The averaged difference electron density map was used for modeling the ligands: F1,6P for the structure of enzyme crystals soaked with F1,6P and DHAP for that soaked with F1P. Unexpected electron density was observed in the active site of the native enzyme structure and modeled as DHAP, on the basis of the HPLC and mass spectrometry analysis described below, and glycerol (GLYC). Water molecules were included in the later rounds of refinement. Final data refinement statistics are listed in Table 2.

Preparation and HPLC Analysis of the O-(4-Nitrobenzyl)-hydroxylamine Derivatives. A total of 100 mg of ADHS at a concentration of 50 mg/mL was used for isolation of the ligand copurified with the protein. The pure protein was buffer exchanged into a 10 mM ammonium acetate solution (pH 6.5), and 150 μL of 0.1 M *O*-(4-nitrobenzyl)hydroxylamine·HCl (NBHA) (Fluka) and 100 μL of 1 M ammonium acetate (pH 4.0) were added. The solution was incubated at 37 °C for 30 min. After the incubation, methanol was added to a final concentration of 30% (v/v). The solution was then heated at 100 °C for 20 min followed by centrifugation (23700*g* for 10 min) to remove the precipitated protein. The clear solute was then decanted from the protein pellet and filtered using a Microcon centrifugal filter device with 10 kDa molecular mass cutoff from Amicon. The solution was lyophilized overnight and dissolved in 500 μL of 10 mM ammonium acetate (pH 6.5) for the HPLC experiments.

F1P (Sigma), fructose 6-phosphate (Fluka), and F1,6P (Sigma) were also derivatized and used for control experiments. The NBHA derivatives for all three hexoses were prepared by adding 150 μL of 0.1 M NBHA and 100 μL of 1 M ammonium acetate (pH 4.0) to 50 μL of a 0.2 M hexose solution. The solutions were then incubated at 37 °C overnight, followed by addition of 200 μL of 10 mM ammonium acetate (pH 6.5) to a final volume of 500 μL. These solutions were stored at −20 °C until HPLC experiments were conducted.

The analysis of the derivatives was performed using a Supelco Discovery BIO Wide Pore C18 column (4.6 mm × 25 mm, 5 μm). All of the samples were eluted with a methanol gradient from 10 to 70% in 10 mM ammonium acetate buffer (pH 6.5), at a flow rate of 0.5 mL/min. The elution profile was monitored at 280 nm.

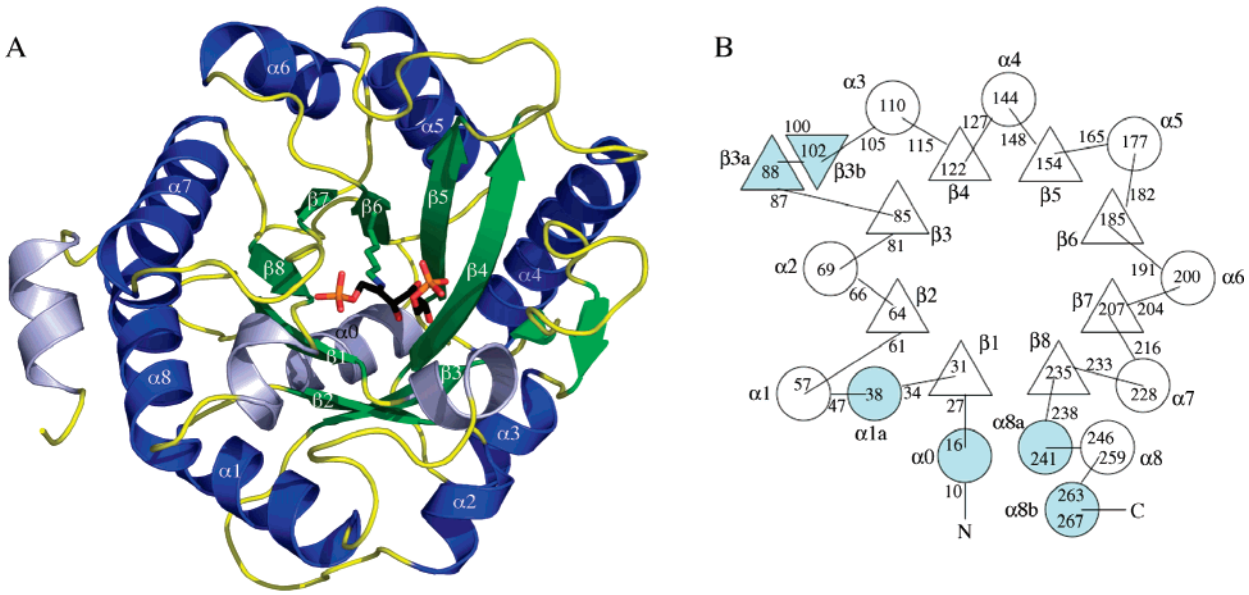


FIGURE 1: Protomer of ADH synthase. (A) A monomer shown as a ribbon diagram. Colored dark green are the strands that are part of the $(\beta\alpha)_8$-barrel, and colored dark blue are the helices that are part of the $(\beta\alpha)_8$-barrel. Colored light green is the pair of antiparallel strands, β3a and β3b. Colored light blue are the additional helices that are not part of the $(\beta\alpha)_8$-barrel fold. In ball-and-stick representation is the ligand, F1,6P. (B) Topology diagram of ADHS. White circles represent helices and white triangles strands of the $(\beta\alpha)_8$-barrel. Numbers inside and outside the circles and triangles refer to the first and last amino acid residues, respectively, for that symmetry element. The light blue helices and strands are the additional secondary elements that are not part of the $(\beta\alpha)_8$-barrel.

ES-MS/MS Analysis of the NBHA Derivatives. Infusion electrospray mass spectrometric analysis was performed on the NBHA derivative of the copurified ADH ligand; NBHA and the NBHA derivative of F1P were used as controls. Prior to MS analysis, all of the samples were diluted to an estimated concentration of 0.500−2 pmol/μL in a solution containing 50% acetonitrile and 0.1% formic acid. Each sample was delivered using a syringe pump to a hybrid triple-quadrupole linear ion trap mass spectrometer, 4000 Q Trap from ABI/MDS Sciex (Framingham, MA), equipped with Turbo V source. The flow rate of the delivery of the sample was 4 μL/min. The sample was analyzed in tune mode with Analyst version 1.4.1. The spray voltage was 5 kV for the positive ion scan mode and −4.5 kV for the negative ion scan mode. Nitrogen was used as the curtain (10 psi) and as a collision gas (set to medium) with a heated interface. The declustering potential was set at 70 eV, and the ion source nebulizer gas was set up to 25 psi. The survey MS scans were acquired in both positive and negative ion Q1 scans (m/z 60−600). To selectively detect the potential phosphate compounds, a precursor ion scan of a marker ion at m/z −79 and a neutral loss scan at m/z +98 were performed and subsequently followed by scans on the ions of interest. All scans were performed for a 5 min acquisition for each sample.

RESULTS

Quality Assessment of the Final Models. Three complexed structures of ADHS were determined in this work. The ADHS−F1,6P complex resulted from soaking the enzyme crystals with F1,6P. The ADHS−DHAP structure resulted from soaking the crystals with F1P, and the ADHS−DHAP−GLYC complex resulted from copurification of the ligands with the native enzyme. The ADHS monomer contains 273 residues, residues 2−272 of which are included in the final models for all of the structures. However, residues 72−79, located in a loop region, exhibited variability between monomers; for most of the monomers, only the backbone of these residues was included in the final models due to the lack of clear side chain electron density, while for some of the monomers, both the backbone and the side chain density were present. The final quality of the models was verified using PROCHECK (*17*). Final refinement statistics, including R factor and R_{free} values, are listed in Table 2.

Overall Structure. The protomer of ADHS has a $(\beta\alpha)_8$-barrel fold (Figure 1). In addition to the core fold, the monomer contains four additional helices and a pair of β-strands. At the N-terminus, an additional helix α0 (residues 10−16) caps the barrel from the bottom. Helix α1a (residues 34−38) inserted between β1 and α1 of the barrel is important for oligomer interface formation. At the C-terminus, α8a (residues 238−241) precedes α8 and is involved in ligand binding. Helix α8b consists of residues 263−267 and follows α8. A pair of antiparallel β-strands (β3a and β3b) found between β3 and α3 of the barrel is important for monomer interaction in the formation of the biological oligomer.

In the crystal structure, ADHS forms a decamer. The decamer consists of two doughnut shaped pentamers with D5 symmetry (Figure 2). Three different monomer−monomer interfaces are formed by the decamer. The A−B interface, between monomers within a pentamer, is formed through hydrogen bonds and hydrophobic contacts of approximately 30 residues from each monomer. In monomer A, helices α4, α5, and α6 and the turns around these helices interact with helices α1, α2, α3, and α4 and the turns associated with those helices from monomer B. The additional helix α1a and the pair of antiparallel β-strands are also part of this interface. A total of 12 300 Å^2, corresponding to 21% of the pentamer surface area, is buried in the interface between the monomers.

Most of the contacts between the two pentamers are between monomers directly stacked (labeled A and C in

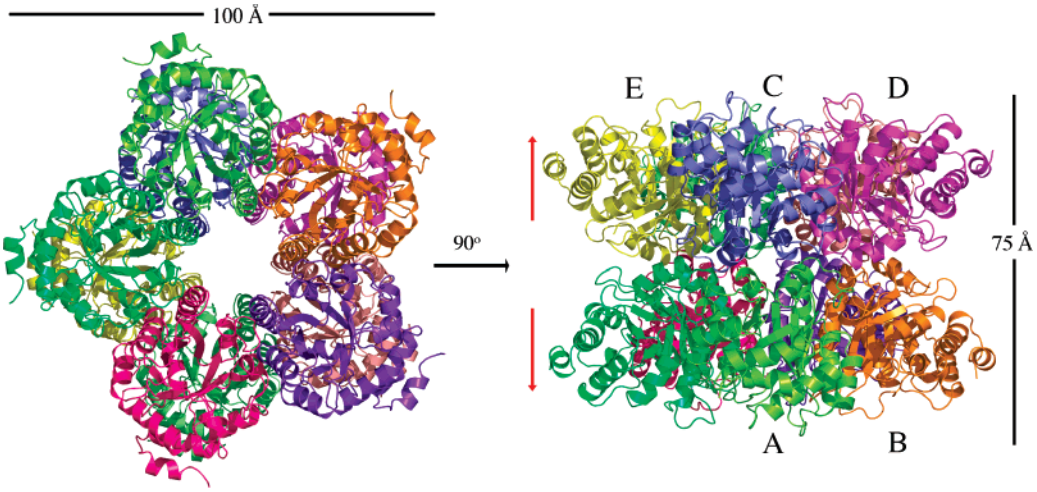


FIGURE 2: Decamer of ADHS. Each monomer is a different color. Two views looking down the 5-fold axis and after a 90° rotation perpendicular to the 5-fold axis. The longest dimensions of the decamer are indicated. The red arrows represent the axis directionality of the $(\beta\alpha)_8$-barrels in the respective pentamers. The monomers labeled A−E are used to describe the interface formation (see the text).

Figure 2), each contributing the same electrostatic interactions between charged residues and hydrophobic contacts. A total of approximately 20 residues participate in the formation of this interface; these residues come from helices α0, α2, α3, and α4 and the turns associated with them. Last, a small monomer A−monomer D interface is formed by five residues, Ile7, Lys8, Leu10, and Leu13 from helix α0 and Tyr143 from helix α4, from both of the monomers, resulting in a mostly hydrophobic environment. There is no interaction between monomers A and E (Figure 2). A total area of 21 800 Å^2 is buried between the two pentamers, comprising 24% of the total decamer surface area. Analytical gel filtration and dynamic light scattering experiments (data not shown) confirm that the decameric state of MjADHS also exists in solution. In contrast, TtFBPA, which was used for molecular replacement, is found as a mixture of pentamers and decamers in solution (*18*). Moreover, in the crystal structure, the decamer interface of TtFBPA makes up only 2% of the total surface area, resulting in a clearly different interface between protomers in this structure (*12*).

Description of the Active Site. The ADHS structure revealed the active site of the enzyme located at the top of the barrel (Figure 1A). The ADHS−F1,6P complex contained the ligand covalently bound to Lys184 located on strand β6 of the barrel (Figure 3A) like that described for TtFBPA (*18*). Lys184 is one of the two strictly conserved lysines in the homologues of this enzyme. The 1-phosphate of the fructose is hydrogen bonded to the side chains of His34 and Arg238, as well as the main chain amide moieties of Arg238, Gly209, and Gly237. The 3- and 5-hydroxy groups of F1,6P make hydrogen bonds with Asp33, while the 4-hydroxy is hydrogen bonded to the side chains of Tyr153 and His119. The 6-phosphate interacts with the side chain of Arg155. All of the residues making contacts with the ligand are strictly conserved among the proteins predicted as ADHSs on the basis of gene clustering and sequence alignment (*11*).

The crystals of ADHS soaked with F1P resulted in a structure containing a clear DHAP electron density linked to Lys184, suggesting that ADHS might possess F1P aldolase activity (Figure 3B). The contacts with 1-phosphate of DHAP are identical to those in the ADHS−F1,6P complex, and the 3-hydroxy of the ligand forms two hydrogen bonds to Asp33. The ADHS−DHAP−GLYC complex structure resulted from the crystals of the native enzyme with a copurified ligand(s) (Figure 3C). The limited resolution of this complex and variability of the ligand occupancy between active sites of the oligomer in the asymmetric unit prevented the identification of the ligand(s) on the basis of the interpretation of the electron density alone. This structure prompted the isolation and identification of the copurified ligand(s) utilizing HPLC and mass spectrometry (described below), subsequently modeled as DHAP and glycerol. DHAP binding here is identical to that in the ADHS−DHAP structure, while glycerol is within hydrogen bonding distance of His65, His85, His119, Ser37, and Lys99, all of which are strictly conserved among ADHSs.

Identification of the Copurified Ligand. The *E. coli* metabolite pool is not known to contain 6-deoxy-5-keto-fructose 1-phosphate (DKFP), the substrate of ADHS forming the Schiff base in the active site. F1,6P and F1P are present in *E. coli*, and F1,6P was shown to bind ADHS in this work; therefore, these sugars could copurify with the native enzyme and thus first were selected for HPLC and MS identification. The NBHA derivative of F1P eluted in 78 min as a single peak, and that of the F1,6P derivative eluted as two peaks at 64 and 68 min. However, the derivative of the copurified ligand(s) eluted as two major peaks at 90 and 100 min. Thus, the copurified ligand(s) was neither of the two hexose-Ps and was less polar then either one.

The fractions corresponding to the two peaks observed by HPLC and containing the copurified ligand(s) derivative(s) were each subjected to the mass spectrometric analysis and yielded identical spectra. The precursor ion scan at *m*/*z* −79 and the neutral loss scan at *m*/*z* +98 gave most abundant ions with masses of *m*/*z* 319 and 321, respectively (Figures 1 and 2 of the Supporting Information). Major fragmentation peaks for the *m*/*z* 319 ion were at *m*/*z* 79, 97, and 136 (listed in order of decreasing intensity) (Figure 3 of the Supporting Information). Major fragmentation peaks for the *m*/*z* 321 ion were at *m*/*z* 223, 168, 88, 303, and 205 (Figure 4 of the Supporting Information). The ion masses, as well as the fragmentation patterns, are consistent with that of a NBHA triose-P derivative and resulted in identification of the

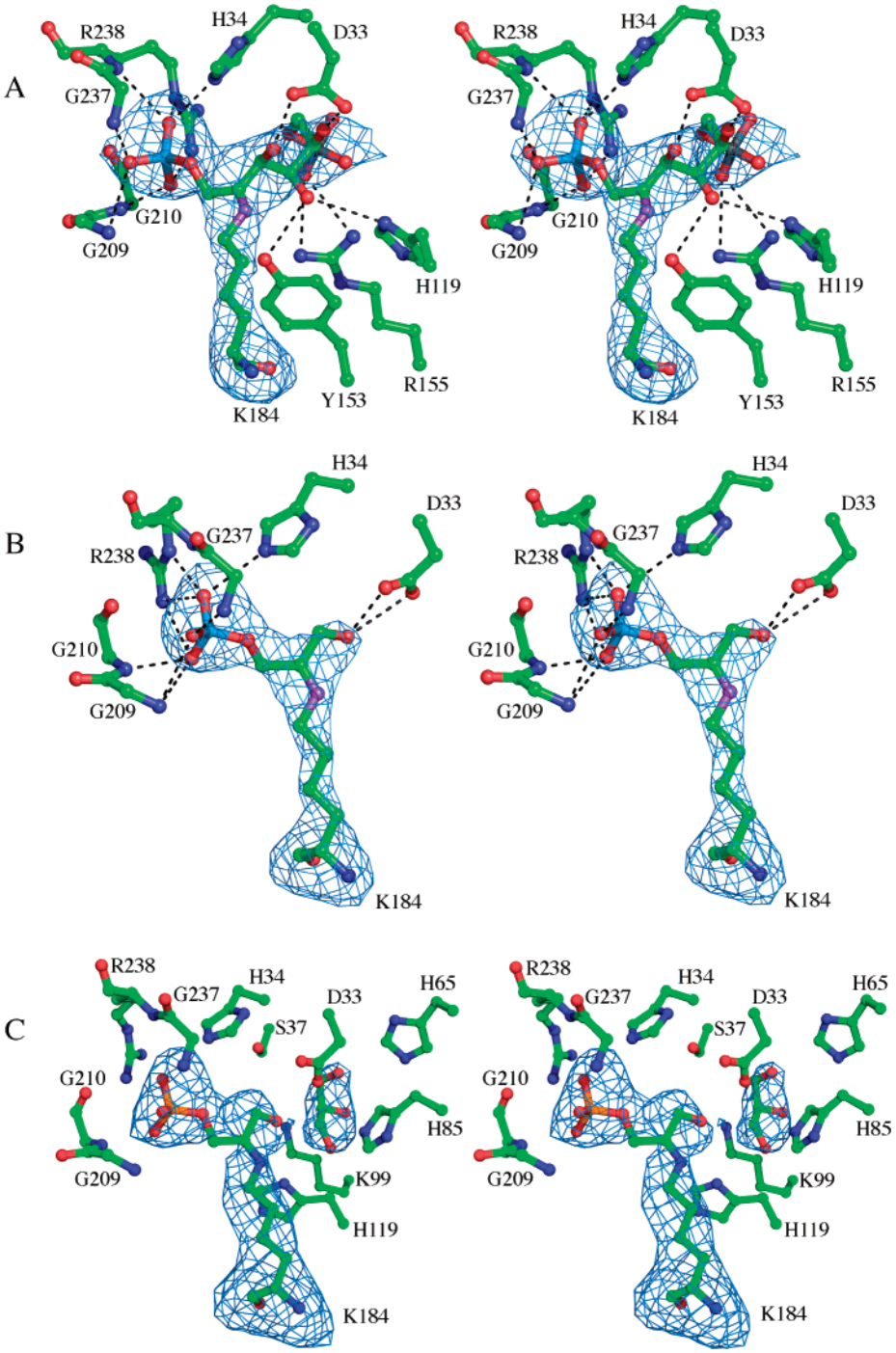


FIGURE 3: Stereoview of the ADHS active site containing different ligands. (A) ADHS−F1,6P complex. The ligand and the residues interacting with the ligand are shown in ball-and-stick representation. Dashed lines are hydrogen bonds. The $F_o - F_c$ averaged electron density, calculated excluding ligand atom phases, is colored blue for the ligand with a contour level of 4σ. (B) Active site for the ADHS−DHAP complex. The presentation is the same as in panel A. (C) Active site of the ADHS−DHAP−GLYC complex.

copurified ligand as the NBHA derivative of DHAP. The observation of two peaks in the HPLC analysis can be explained by the presence of the *syn* and *anti* isomers of the oximine.

However, DHAP alone is a poor fit to the experimental electron density in the active site of the structure containing the copurified ligand(s). The size and shape of the density could reasonably accommodate a glycerol molecule along with DHAP. Glycerol was present at high concentrations for cryoprotection and is likely to bind as an analogue for glyceraldehyde, the other product of F1P cleavage along with DHAP. Thus, combining the results from the crystallographic, HPLC, and MS experiments, we generated a model of ADHS in complex with DHAP and glycerol.

DISCUSSION

ADHS Is a Class IA Aldolase. The sequence of MjADHS is 35% identical with that of TtFBPA. Both are homodecamers containing the active site in the top of the barrel and form a covalent adduct with a substrate utilizing a strictly conserved lysine residue located on strand β6 of the barrel. Also, like the archaeal FBPA, ADHS contains a tyrosine

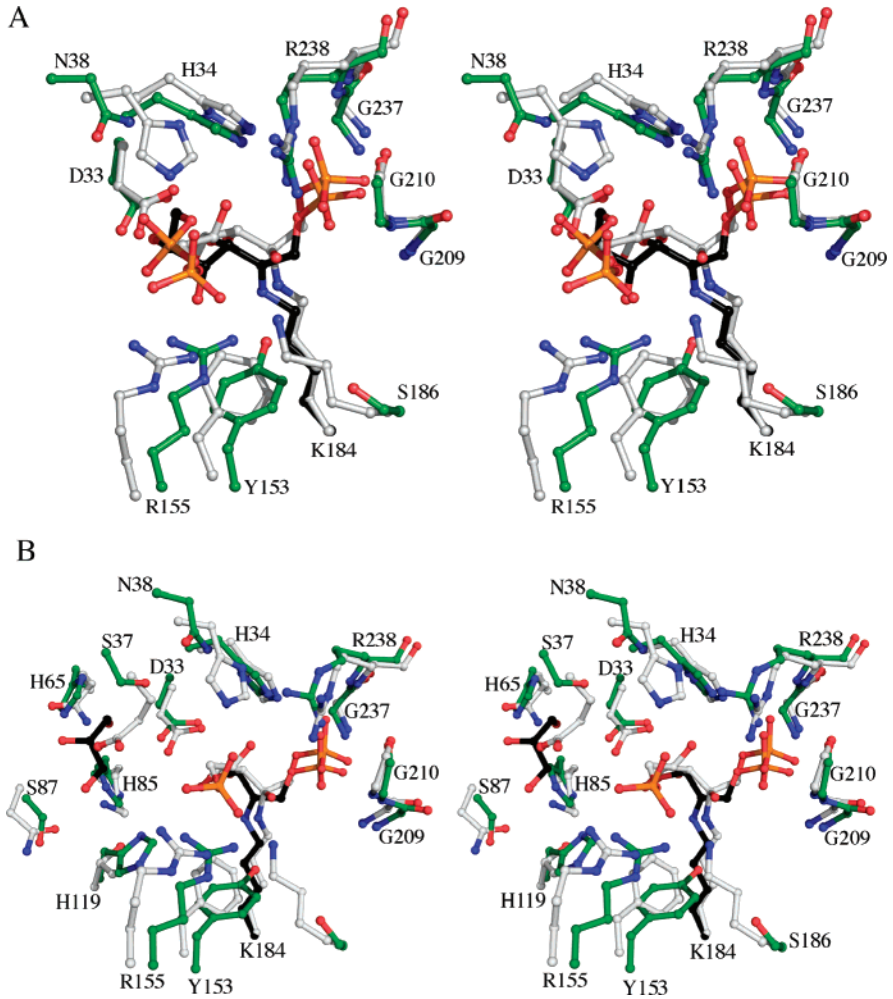


FIGURE 4: Comparison of the active sites of MjADHS and TtFBPA. (A) Superposition of the ADHS−F1,6P complex with the TtFBPA-Y146F−F1,6P complex. Colored green are the residues of MjADHS, and colored black is the ligand of MjADHS covalently attached to Lys184; colored gray are the residues from TtFBPA. The labels are for the residues from MjADHS. (B) Superposition of the TtFBPA native structure with the MjADHS−DHAP−GLYC complex. Color coding and labels are the same as in panel A.

Table 3: Structural Comparison of ADHS with Protein Structures Based on a DALI Search

protein	PDB entry	Z^a	rmsd[b] (Å)	LALI[c]	% IDE[d]
T. tenax fructose 1,6-bisphosphate aldolase	1OJX	34.1	1.7	242	35
S. typhimurium protein of unknown function	1TO3	20.6	2.6	220	17
P. woessei triose phosphate isomerase	1HG3	17.1	2.3	180	12
rabbit muscle aldolase	1ZAH	16.9	3.1	222	13

[a] The *Z* score is the strength of structural similarity in standard deviations above expected. [b] Positional root-mean-square deviation of superimposed Cα atoms. [c] Total number of equivalent residues. [d] Percentage of sequence identity over equivalent residues.

residue located on strand β5, which is part of the catalytic dyad responsible for formation of the Schiff base in FBPA as well as the other members of the group of enzymes using Schiff bases as intermediates (*19*). The fold of the enzyme, the utilization of a Schiff base intermediate, and the presence and location of the catalytic Lys-Tyr dyad are all characteristic of the class I aldolase/transaldolase superfamily, therefore confirming ADHS as a new member of this superfamily (*20*−*23*).

The DALI search based on an ADHS monomer revealed four additional structures closely related to this enzyme (Table 3) (*24*). These structures are of enzymes with known or predicted aldolase activity, FBPA aldolase from rabbit muscle (RAMA) and triose phosphate isomerase from *Pyrococcus woesei* (PwTIM), which bind the same ligands as ADHS. The active sites for these enzymes are also located at the top of the barrel, and they all contain a conserved phosphate binding pocket which corresponds to the binding pocket of the 1-phosphate of F1,6P in MjADHS.

Comparison of the Active Sites of MjADHS and TtFBPA. The superposition of the ADHS−F1,6P complex with the TtFBPA-Y146F−F1,6P structure (PDB entry 1W8R) reveals a similar binding of the ligand (Figure 4A) (*19*). Structure-based sequence alignment of these two enzymes reveals that approximately 60% of the residues are conserved between the two active sites (Figure 5). The 1-phosphate binding pocket of F1,6P, consisting of strictly conserved residues 237−239 in MjADHS (a GRN motif), Gly209, Gly210, and His34, is present in both structures. Also, residues Asp33 and Tyr153, ligands to the 3-, 4-, and 5-hydroxyls of the sugar and catalytically important in TtFBPA, are structurally conserved in ADHS. One of the important differences in the

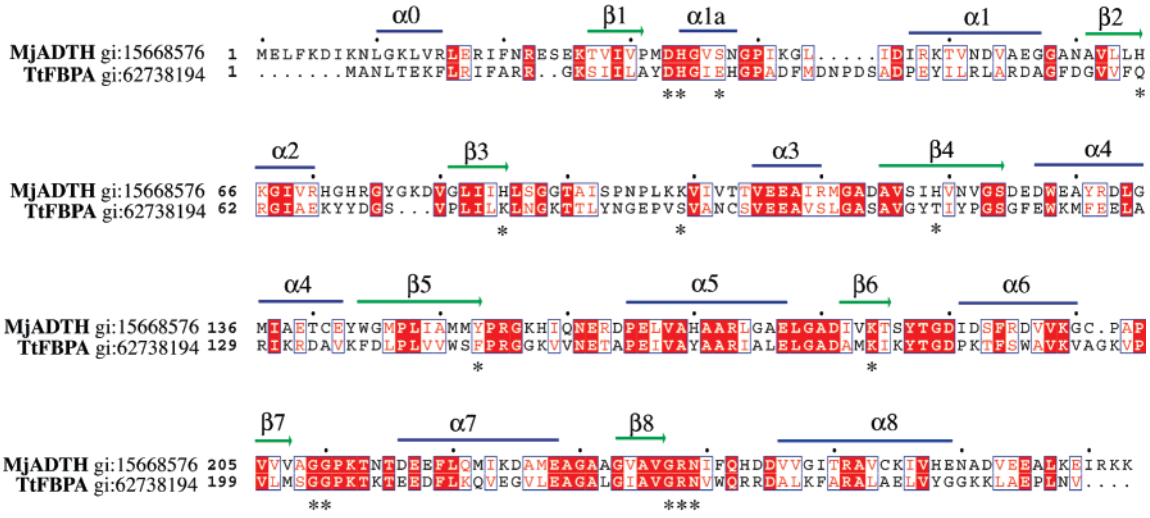
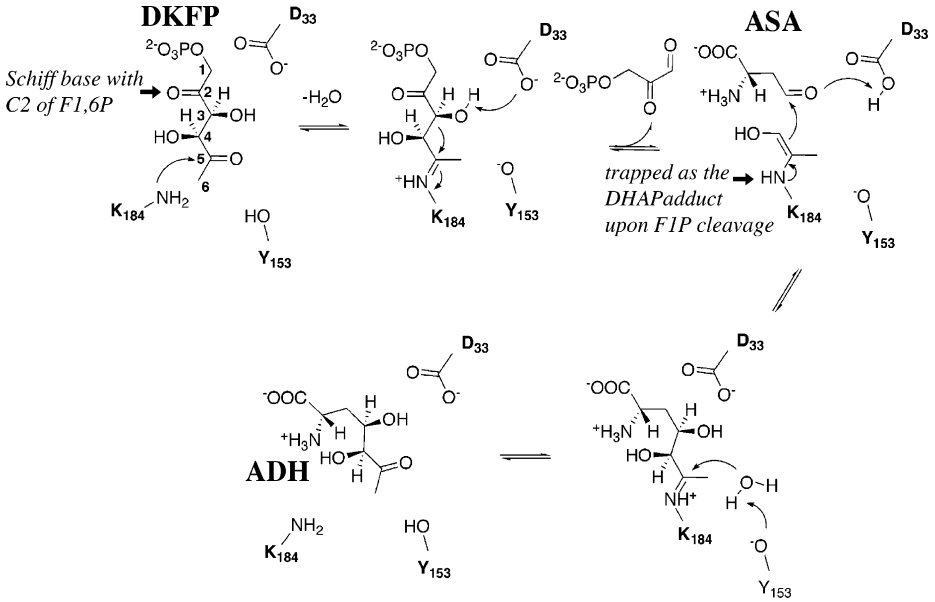


FIGURE 5: Structure-based sequence alignment of MjADHS and TtFBPA. Strictly conserved residues are highlighted in red, and strongly conserved residues are colored red. Secondary structure elements of MjADHS are labeled at the top of the alignment; active site residues in the *M. jannaschii* structure are labeled with an asterisk.

Scheme 2: Proposed Mechanism for ADH Formation



ligand binding sites between the two structures is the 6-phosphate binding pocket. Here R155 is conserved, but the other positively charged residues present in TtFBPA, His29 and Lys179, are replaced with polar residues Asn38 and Ser186. It is likely that this region of the active site is responsible for some of the substrate specificity of ADHS. Another significant difference is revealed upon the superposition of TtFBPA with the ADHS−DHAP−GLYC structure. The strictly conserved active site pocket accommodating glycerol in ADHS is not conserved in the FBPA structure (Figure 4B). Thus, this part of the active site, unique to ADHS, is also likely to be important for substrate specificity and catalysis.

Implications for the Catalytic Mechanism. The availability of the complexed structures reported herein provides insight into the location of the active site of ADHS, as well as the description of the key residues involved in ligand binding and catalysis. On the basis of this information and comparison with FBPA, ADHS is likely to catalyze the cleavage of DKFP utilizing a class I aldolase mechanism (*25*) (Scheme 2). Analogously with the archaeal FBPA catalytic mechanism in ADHS, Tyr153 should be the general acid protonating the C5-hydroxyl of the carbinolamine intermediate and activating it as a leaving group in the formation of the Schiff base in the forward reaction. Then, Asp33 is the general base responsible for the deprotonation of the C3-hydroxyl moiety and facilitation of cleavage of the C3−C4 bond (*26*). An analogous mechanism has been proposed for other members of class I aldolases, among them 2-keto-3-deoxy-6-phosphogluconate aldolase and transaldolase B (*27*, *28*).

The physiological substrate of ADHS, DKFP, contains two carbonyl moieties capable of Schiff base formation. If the mechanism is to proceed via the proposed aldol condensation reaction, the C5-carbonyl should be the one forming the covalent adduct, resulting in the availability of the Schiff base electron sink and the formation of an eneamine capable of attacking L-aspartate semialdehyde (ASA). A model of DKFP bound via a C5-Schiff base in the active site of MjADHS based on the ADHS−F1,6P structure reveals that the proposed key catalytic residues are within a reasonable

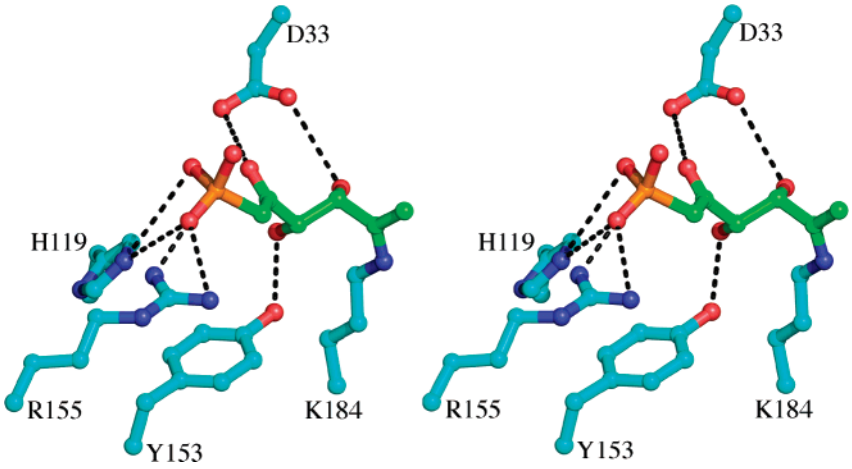


FIGURE 6: Model of the DKFP Schiff base intermediate in the active site of MjADHS. Residues within hydrogen bonding distance of the ligand are shown.

distance required for catalysis (Figure 6). However, the strict conservation of the 1-phosphate binding site as well as the intricate hydrogen bonding network associated with this phosphate binding is observed, while no phosphate binding pocket is apparent at the C6 position from the ADH−F1,6P complex. Thus, structurally, formation of the Schiff base with the C2-carbonyl moiety would result in a favorable binding mode for the 1-phosphate group. The ability to utilize the C2-carbonyl is further supported by the ADHS−DHAP complex, where the observation of the product of F1P cleavage is still consistent with the mechanism proposed in Scheme 2, the difference being in the location of the reactive carbonyl that is used. Thus, it appears that ADHS could display aldolase activity with either of the two carbonyl moieties. Should the aldolase cleavage occur from the C2 Schiff base-bound DKFP, the products of the reaction would be DHAP and methylglyoxal, the latter of which would be converted back into DKFP, by the gene product of Mj1585, another enzyme designated as an aldolase in *M. jannaschii* (*29*).

Substrate Specificity. ADHS does not cleave F1,6P as reported for the *Methanococcus maripaludis* homologue (*11*). However, it binds this ligand as seen in the structure of the ADHS−F1,6P complex. Despite soaking conditions identical to those for F1,6P, F1P was not observed in the active site; instead, a DHAP molecule was bound. The ADHS−DHAP complex is suggestive of F1P aldolase activity. A broad range of substrate specificity has previously been observed in aldolases and also appears to be generally more characteristic in enzymes of archaeal organisms (*30*−*32*). The broad specificity would be in accord with proposed various binding modes of the ligands, and formation of the Schiff base with either a C2- or C5-carbonyl as suggested above.

Product release is a rate-limiting step in the reaction of FBPA and also type I dehydroquinase, another member of the class I aldolase superfamily, hence supporting the observation of the covalently bound product, DHAP, in the ADHS structures (*22*). Moreover, since F1P is not the physiological substrate, the products could be trapped in the crystal due to nonphysiological rates of cleavage and release. Another possibility is that ASA has to be present for the reaction to proceed forward with the formation of ADH, which in turn could trigger the release of the products. Efforts to obtain a structure with ASA or its analogues so far have been unsuccessful, and the binding location of ASA is unknown. The active site pocket is large and capable of accommodating ASA in several binding modes, as well as both substrates simultaneously, making modeling studies challenging. Thus, the order of addition and release of the substrates and products cannot be deduced from the structure alone. While the complexes described herein provide initial insight into the possible roles of the active site residues, biochemical studies of the wild-type enzyme and analyses of site-directed mutants are essential for addressing the mechanistic details definitively. These studies should also reveal the evolutionary connection between the two related transaldolases present in *M. jannaschii*: the Mj0400 gene product described here, which uses ASA and DKFP as substrates, and the multifunctional Mj1585 gene product, which is an FBP aldolase (*8*) and also catalyzes the formation of DKFP (*29*).

Conclusions. The ADHS structure reported in this work identifies this enzyme as a member of the class I aldolase/transaldolase superfamily. The ADHS complexes with F1,6P, DHAP, and DHAP and glycerol together enable the characterization of the active site of the enzyme. On the basis of these structures, functions of several residues that may be catalytically important are predicted: Lys184 forms the Schiff base intermediate, and Tyr153 and Asp33 are likely the general acids/bases aiding Schiff base formation and substrate cleavage. Several additional strictly conserved residues, His65, His85, Lys99, and His119, were also identified in the active site; however, their function in the catalytic mechanism remains unclear and requires further characterization.

ACKNOWLEDGMENT

We thank Leslie Kinsland for assistance in the preparation of the manuscript. We thank the NE-CAT staff at beamline 24-ID-C of the Advanced Photon Source for assistance with data collection. We thank Robert Gahl for the assistance with the HPLC experiments. The Protein Facility of the Department of Chemistry and Chemical Biology and the Proteomics and Mass Spectrometry facility of the Biotechnology Resource Center at Cornell University are gratefully acknowledged.

SUPPORTING INFORMATION AVAILABLE

Negative precursor *m*/*z* −79 ion scan for the NBHA derivative of the copurified ligand (Figure 1), fragmentation pattern of the *m*/*z* 319 ion selected from the precursor *m*/*z* −79 ion scan (Figure 2), neutral loss scan at *m*/*z* +98 for the NBHA derivative of the copurified ligand (Figure 3), and fragmentation pattern of the *m*/*z* 321 ion selected from the neutral loss scan at *m*/*z* +98 (Figure 4). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Pittard, A. J. (1996) *Escherichia coli* and *Salmonella typhimurium* cellular and molecular biology, in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., Eds.) pp 458−484, American Society for Microbiology, Washington, DC.
2. Shumilin, I. A., Kretsinger, R. H., and Bauerle, R. H. (1999) Crystal structure of phenylalanine-regulated 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*, *Structure 7*, 865−875.
3. Carpenter, E. P., Hawkins, A. R., Frost, J. W., and Brown, K. A. (1998) Structure of dehydroquinate synthase reveals an active site capable of multistep catalysis, *Nature 394*, 299−302.
4. Bohlke, K., Pisani, F. M., Rossi, M., and Antranikian, G. (2002) Archaeal DNA replication: Spotlight on a rapidly moving field, *Extremophiles 6*, 1−14.
5. White, M. F. (2003) Archaeal DNA repair: Paradigms and puzzles, *Biochem. Soc. Trans. 31*, 690−693.
6. Daugherty, M., Vonstein, V., Overbeek, R., and Osterman, A. (2001) Archaeal shikimate kinase, a new member of the GHMP-kinase family, *J. Bacteriol. 183*, 292−300.
7. Hochuli, M., Patzelt, H., Oesterhelt, D., Wuthrich, K., and Szyperski, T. (1999) Amino acid biosynthesis in the halophilic archaeon *Haloarcula hispanica*, *J. Bacteriol. 181*, 3226−3237.
8. Makarova, K. S., Aravind, L., Galperin, M. Y., Grishin, N. V., Tatusov, R. L., Wolf, Y. I., and Koonin, E. V. (1999) Comparative genomics of the Archaea (*Euryarchaeota*): Evolution of conserved protein families, the stable core, and the variable shell, *Genome Res. 9*, 608−628.
9. Porat, I., Waters, B. W., Teng, Q., and Whitman, W. B. (2004) Two biosynthetic pathways for aromatic amino acids in the archaeon *Methanococcus maripaludis*, *J. Bacteriol. 186*, 4940−4950.
10. White, R. H. (2004) L-Aspartate semialdehyde and a 6-deoxy-5-ketohexose 1-phosphate are the precursors to the aromatic amino acids in *Methanocaldococcus jannaschii*, *Biochemistry 43*, 7618−7627.
11. Porat, I., Sieprawska-Lupa, M., Teng, Q., Bohanon, F. J., White, R. H., and Whitman, W. B. (2006) Biochemical and genetic characterization of an early step in a novel pathway for the biosynthesis of aromatic amino acids and *p*-aminobenzoic acid in the archaeon *Methanococcus maripaludis*, *Mol. Microbiol. 62*, 1117−1131.
12. Lorentzen, E., Pohl, E., Zwart, P., Stark, A., Russell, R. B., Knura, T., Hensel, R., and Siebers, B. (2003) Crystal structure of an archaeal class I aldolase and the evolution of $(\beta\alpha)_8$ barrel proteins, *J. Biol. Chem. 278*, 47253−47260.
13. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
14. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR System: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
15. Kleywegt, G. J., and Jones, T. A. (1996) xdlMAPMAN and xdlDATAMAN: Programs for reformatting, analysis, and manipulation of biomacromolecular electron-density maps and reflection datasets, *Acta Crystallogr. D52*, 826−828.
16. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics, *Acta Crystallogr. D60*, 2126−2132.
17. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
18. Siebers, B., Brinkmann, H., Dorr, C., Tjaden, B., Lilie, H., van der Oost, J., and Verhees, C. H. (2001) Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase, *J. Biol. Chem. 276*, 28710−28718.
19. Lorentzen, E., Siebers, B., Hensel, R., and Pohl, E. (2005) Mechanism of the Schiff base forming fructose-1,6-bisphosphate aldolase: Structural analysis of reaction intermediates, *Biochemistry 44*, 4222−4229.
20. Choi, K. H., Lai, V., Foster, C. E., Morris, A. J., Tolan, D. R., and Allen, K. N. (2006) New superfamily members identified for Schiff-base enzymes based on verification of catalytically essential residues, *Biochemistry 45*, 8546−8555.
21. Babbitt, P. C., and Gerlt, J. A. (1997) Understanding enzyme superfamilies. Chemistry as the fundamental determinant in the evolution of new catalytic activities, *J. Biol. Chem. 272*, 30591−30594.
22. Lawrence, M. C., Barbosa, J. A., Smith, B. J., Hall, N. E., Pilling, P. A., Ooi, H. C., and Marcuccio, S. M. (1997) Structure and mechanism of a sub-family of enzymes related to N-acetylneuraminate lyase, *J. Mol. Biol. 266*, 381−399.
23. Leech, A. P., James, R., Coggins, J. R., and Kleanthous, C. (1995) Mutagenesis of active site residues in type I dehydroquinase from *Escherichia coli*. Stalled catalysis in a histidine to alanine mutant, *J. Biol. Chem. 270*, 25827−25836.
24. Holm, L., and Sander, C. (1998) Touring protein fold space with Dali/FSSP, *Nucleic Acids Res. 26*, 316−319.
25. Gefflaut, T., Blonski, C., Perie, J., and Willson, M. (1995) Class I aldolases: Substrate specificity, mechanism, inhibitors and structural aspects, *Prog. Biophys. Mol. Biol. 63*, 301−340.
26. Choi, K. H., Shi, J., Hopkins, C. E., Tolan, D. R., and Allen, K. N. (2001) Snapshots of catalysis: The structure of fructose-1,6-(bis)phosphate aldolase covalently bound to the substrate dihydroxyacetone phosphate, *Biochemistry 40*, 13868−13875.
27. Allard, J., Grochulski, P., and Sygusch, J. (2001) Covalent intermediate trapped in 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase structure at 1.95 Å resolution, *Proc. Natl. Acad. Sci. U.S.A. 98*, 3679−3684.
28. Schorken, U., Thorell, S., Schurmann, M., Jia, J., Sprenger, G. A., and Schneider, G. (2001) Identification of catalytically important residues in the active site of *Escherichia coli* transaldolase, *Eur. J. Biochem. 268*, 2408−2415.
29. White, R. H., and Xu, H. (2006) Methylglyoxal is an intermediate in the biosynthesis of 6-deoxy-5-ketofructose-1-phosphate: A precursor for aromatic amino acid biosynthesis in *Methanocaldococcus jannaschii*, *Biochemistry 45*, 12366−12379.
30. Subramaniam, P. S., Xie, G., Xia, T., and Jensen, R. A. (1998) Substrate ambiguity of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* in the context of its membership in a protein family containing a subset of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthases, *J. Bacteriol. 180*, 119−127.
31. Schofield, L. R., Anderson, B. F., Patchett, M. L., Norris, G. E., Jameson, G. B., and Parker, E. J. (2005) Substrate ambiguity and crystal structure of *Pyrococcus furiosus* 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase: An ancestral 3-deoxyald-2-ulosonate-phosphate synthase? *Biochemistry 44*, 11950−11962.
32. Jensen, R. A. (1976) Enzyme recruitment in evolution of new function, *Annu. Rev. Microbiol. 30*, 409−425.

BI700934V