THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 21, Issue of May 23, pp. 19463–19472, 2003

Printed in U.S.A.

Structures of Shikimate Dehydrogenase AroE and Its Paralog YdiB

A COMMON STRUCTURAL FRAMEWORK FOR DIFFERENT ACTIVITIES*

Received for publication, January 24, 2003, and in revised form, March 10, 2003
Published, JBC Papers in Press, March 12, 2003, DOI 10.1074/jbc.M300794200

Gurvan Michel‡, Aleksander W. Roszak§¶, Véronique Sauvé‡, John Maclean§, Allan Matte‡, John R. Coggins¶, Miroslaw Cygler‡‖, and Adrian J. Lapthorn§**

From the ‡Biotechnology Research Institute, NRC Macromolecular Structure Group, Montreal, Quebec H4P 2R2, Canada and the §Department of Chemistry and ¶Division of Biochemistry and Molecular Biology, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow G12 8QQ, Scotland, United Kingdom

Shikimate dehydrogenase catalyzes the fourth step of the shikimate pathway, the essential route for the biosynthesis of aromatic compounds in plants and microorganisms. Absent in metazoans, this pathway is an attractive target for nontoxic herbicides and drugs. *Escherichia coli* expresses two shikimate dehydrogenase paralogs, the NADP-specific AroE and a putative enzyme YdiB. Here we characterize YdiB as a dual specificity quinate/shikimate dehydrogenase that utilizes either NAD or NADP as a cofactor. Structures of AroE and YdiB with bound cofactors were determined at 1.5 and 2.5 Å resolution, respectively. Both enzymes display a similar architecture with two α/β domains separated by a wide cleft. Comparison of their dinucleotide-binding domains reveals the molecular basis for cofactor specificity. Independent molecules display conformational flexibility suggesting that a switch between open and closed conformations occurs upon substrate binding. Sequence analysis and structural comparison led us to propose the catalytic machinery and a model for 3-dehydroshikimate recognition. Furthermore, we discuss the evolutionary and metabolic implications of the presence of two shikimate dehydrogenases in *E. coli* and other organisms.

The shikimate pathway, which links metabolism of carbohydrates to biosynthesis of aromatic compounds, is essential to plants, bacteria, and fungi (1) as well as apicomplexan parasites (2). This seven-step metabolic route leads from phosphoenolpyruvate and erythrose 4-phosphate to chorismate, the common precursor for the synthesis of folic acid, ubiquinone, vitamins E and K, and aromatic amino acids (1). This pathway is absent in metazoans, which must obtain the essential amino acids phenylalanine and tryptophan from their diet. Therefore, enzymes of this pathway are important targets for the development of nontoxic herbicides (3), as well as antimicrobial (4) and antiparasite (2) agents. The sixth step in the pathway, catalyzed by 5-enolpyruvylshikimate-3-phosphate synthase, has already been successfully targeted, with the development of glyphosate, a broad spectrum herbicide (5). However, after 20 years of extensive use, glyphosate-resistant weeds have recently emerged (6), emphasizing the importance of maintaining target diversity. In order to design new inhibitors, crystal structures of several enzymes of the shikimate pathway have been elucidated recently: 3-dehydroquinate synthase (7), type I and II dehydroquinases (8), type I and II shikimate kinases (9, 10), and 5-enolpyruvylshikimate-3-phosphate synthase (11), catalyzing the second, third, fifth, and sixth steps of the pathway, respectively.

Shikimate dehydrogenase (EC 1.1.1.25) catalyzes the fourth reaction in the shikimate pathway, the NADP-dependent reduction of 3-dehydroshikimate to shikimate (Fig. 1*A*). Whereas dehydrogenases usually form oligomers, shikimate dehydrogenase, coded by the gene *aroE* in *Escherichia coli*, is present as a monomer in most bacteria (12, 13). In higher organisms this activity is part of a multifunctional enzyme. In plants shikimate dehydrogenase is associated with type I dehydroquinase to form a bifunctional enzyme (14), whereas in fungi, such as *Neurospora crassa*, this enzyme forms the fifth domain of the pentafunctional AROM polypeptide, which catalyzes five of seven steps of the shikimate pathway (15). However, the molecular basis of 3-dehydroshikimate recognition and enzymatic reduction is not known.

Although in *E. coli* AroE is strictly specific for shikimate, some fungal shikimate dehydrogenases can also utilize quinic acid as a substrate. This compound, which differs from shikimic acid only by the addition of a hydroxyl group at C-1 (Fig. 1*B*), is the precursor to the ubiquitous plant secondary product chlorogenate (1). To date, two independent families of quinate/shikimate dehydrogenases have been identified. The first consists of NAD-dependent dehydrogenases (16), and the second consists of membrane-associated dehydrogenases that utilize pyrrolo-quinoline-quinone as a cofactor (17). Both types of dehydrogenases are involved in the catabolic quinate pathway, which allows growth of microorganisms with quinate as the sole carbon source by its conversion into protocatechuate and subsequent metabolism by the β-ketoadipate pathway (16, 17).

By using BLAST (18), ~130 sequences, mostly annotated as putative shikimate dehydrogenases, can be identified as homologous to AroE through the entire length of the gene, thereby defining the shikimate dehydrogenase (SDH)[1] family. It also includes the NAD-dependent quinate/shikimate dehydroge-

* This work was supported in part by Canadian Institutes of Heath Research Grant 200103GSP-90094-GMX-CFAA-19924 (to M. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1NYT and 1O9B) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‖ To whom correspondence may be addressed: Biotechnology Research Institute, NRC Macromolecular Structure Group, 6100 Royalmount Ave., Montreal, Quebec H4P 2R2, Canada. E-mail: mirek@bri.nrc.ca.

** To whom correspondence may be addressed: Dept. of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland, UK. E-mail: adrian@chem.gla.ac.uk.

[1] The abbreviations used are: SDH, shikimate dehydrogenase; DTT, dithiothreitol; r.m.s.d., root mean square deviation.

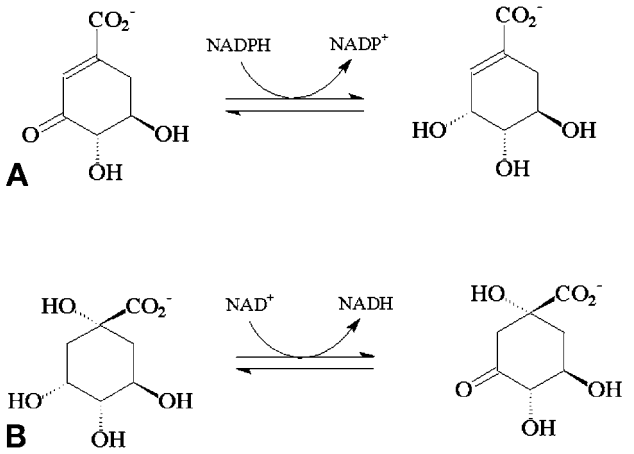

FIG. 1. **The reactions catalyzed by shikimate (*A*) and quinate (*B*) dehydrogenases.**

nases, whereas the pyrrolo-quinoline-quinone-dependent enzymes compose a different protein family. This family displays no significant sequence similarity with any other NAD(P)-dependent dehydrogenases, therefore constituting a distinct dehydrogenase family. Analysis of the complete genome of *E. coli* K12 and pathogenic O157:H7 strains has revealed the presence of a gene of unknown function, *ydiB*, which shares 25% sequence identity with *aroE*. Thus, AroE and YdiB are paralogs, the only two proteins from the SDH family present in *E. coli*.

Here we report the biochemical characterization of YdiB and demonstrate that it is a quinate/shikimate dehydrogenase that can utilize either NAD or NADP as a cofactor. We have determined crystal structures of both these enzymes, AroE at 1.5 Å and YdiB at 2.5 Å resolution. These structures are the first shikimate dehydrogenase structures to be determined. Comparison of their substrate-binding sites led us to propose the catalytically important amino acid residues and to identify at the molecular level the structural differences leading to variation in cofactor specificity. Furthermore, we discuss the evolutionary and metabolic implications of the presence of two shikimate dehydrogenases in *E. coli* and other organisms.

EXPERIMENTAL PROCEDURES

Expression, Purification, and Crystallization—AroE was overexpressed, purified, and crystallized as described previously (19). The crystals belong to space group C-2, with unit cell dimensions a = 110.0 Å, b = 139.5 Å, and c = 102.6 Å, β = 122.2° with four molecules in the asymmetric unit. The *ydib* gene was cloned into a derivative of pGEX-4T1 (Amersham Biosciences) amplified from *E. coli* strain MC1061 genomic DNA using recombinant *Taq* polymerase (Amersham Biosciences). Plasmid DNA was transformed into *E. coli* DL41, and cells were grown in defined LeMaster medium (20) supplemented with 25 mg ml^{-1} L-selenomethionine, at 37 °C to an A_{600} of ~0.8. A 1-liter culture was induced with 100 μM isopropyl-1-thio-β-D-galactopyranoside and culturing continued at 20 °C for 15 h. Cells were harvested by centrifugation and resuspended in 40 ml of lysis buffer (50 mM Tris-HCl, pH 7.5, 400 mM NaCl, 1% (w/v) Triton X-100, 5% (w/v) glycerol, 10 mM DTT, and one tablet of Complete™ protease inhibitors (Roche Diagnostics)). Cells were lysed by sonication, and the lysate was clarified by ultracentrifugation (20,000 × *g*, 4 °C, 30 min). The protein supernatant was applied to a 5-ml (bed volume) DEAE-Sepharose (Amersham Biosciences) column equilibrated with lysis buffer, and the flow-through was collected. This fraction was then applied to a 5-ml column of glutathione-Sepharose 4B (Amersham Biosciences), and the column was washed with 10 bed volumes of buffer (50 mM Tris-HCl, pH 7.5, 1 M NaCl, 1% (w/v), 5% (w/v) glycerol, 10 mM DTT). The target protein was isolated by cleaving the glutathione *S*-transferase fusion while bound to glutathione-Sepharose by addition of human α-thrombin (ratio 1:2000 (w/w), 4 °C, 15 h), followed by elution with buffer (50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1% (w/v), 5% (w/v) glycerol, 10 mM DTT). For crystallization trials, the protein was concentrated by ultrafiltration to 7.0 mg/ml and the buffer changed to 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 5% (w/v) glycerol, 5 mM DTT. Initial crystals were obtained by sparse-matrix screening (Hampton Research, Laguna Niguel, CA) by hanging drop vapor diffusion. The best crystals grew at 21 °C from droplets containing 2 μl of protein solution and 2 μl of reservoir solution (0.8 M KH_2PO_4, 0.8 M NaH_2PO_4, 0.1 M Hepes, pH 7.5, 2 mM NADH) and reached their maximal size after 2 days. The crystals are hexagonal, space group $P6_4$ with cell dimensions $a = b$ = 157.87 Å, c = 40.01 Å, $\alpha = \beta = 90°$, $\gamma = 120°$ with two molecules in the asymmetric unit.

Structure Solution and Refinement—Native data were collected from cryo-cooled AroE crystals to 1.5 Å resolution at station 9.6 at the Daresbury SRS using a (ADSC) Quantum-4 charge-coupled device detector. Data were collected in-house, by using a MacScience DIP2000 detector on various crystals soaked with heavy atom solutions. Crystals soaked with $Hg(CN)_2$ produced the only usable derivative that was isomorphous to the native crystals. This derivative was collected at a wavelength to maximize the anomalous signal on station 9.5 at the Daresbury SRS using a Mar charge-coupled device detector in a SIRAS experiment. All data were indexed and processed with the HKL suite (21), the cell dimensions and space group shown in Table I. Further processing was carried out using programs from the CCP4 package (22). From the anomalous Patterson map it was possible to identify 13 mercury sites using SHELX-90 (23), which were refined in Mlphare against the native 1.5-Å data to maximize the isomorphous signal. Phase refinement and extension was performed using the program DM with solvent flattening and histogram matching. Averaging was attempted but was unsuccessful because of the large variation in conformation in the independent molecules in the asymmetric unit. Refinement was carried out using the maximum likelihood refinement program REFMAC (24). Five percent of the data were randomly set aside as test data for calculation of R_{free}. The structure was built automatically using the program ARP/WARP (25) and assembled into the four independent chains that were >90% complete. Manual correction of the structure and model building and addition of solvent was performed using modules within the program QUANTA (Accelrys Inc.). Nine iterations of refinement and manual rebuilding with the addition of molecules of $NADP^+$, sulfate, glycerol, and DTT with the application of individual anisotropic temperature factors in the final stages of refinement resulted in a model with the final R_{work} of 14.7% and R_{free} of 17.6% and good stereochemistry as assessed using the program PROCHECK (26). The structure was deposited with the Protein Data Bank with the code 1NYT.

YdiB crystals were soaked for ~30 s in a cryoprotectant solution (0.8 M KH_2PO_4, 0.8 M NaH_2PO_4, 0.1 M Hepes, pH 7.5, 2 mM NADH, 22% (v/v) glycerol), picked up in a nylon loop, transferred to the goniometer head, and kept at 100 K in a nitrogen stream. Diffraction data were collected on a Quantum-4 charge-coupled device detector (ADSC, San Diego, CA) at beamline X8C, the National Synchroton Light Source (NSLS) at Brookhaven National Laboratory in New York. Data indexing, merging, and scaling were performed using the HKL2000 package (21). Data collection and processing statistics are listed in Table I. Multiple anomalous dispersion data were collected on a Se-Met-substituted YdiB crystal to 2.5 Å resolution at inflection, peak, and hard remote wavelengths around the K absorption edge of selenium (Table I). Of the 22 expected selenium sites, 20 were found using the heavy atom search procedure of CNS (27). The phases calculated with this partial structure resulted in a figure of merit of 0.67–2.5 Å resolution. By taking advantage of the noncrystallographic symmetry (NCS), the electron density was improved by molecular averaging and solvent flipping (40% solvent) with CNS (27), yielding a final figure of merit of 0.93. The model was built manually with the program O (28) into the solvent-flipped multiple anomalous dispersion electron density map. Refinement was performed with CNS (27) with the maximum likelihood target function. The NCS restraints were applied only in the initial cycles of refinement. The experimental as well as the simulated annealing omit maps clearly showed the presence of one NAD^+ molecule bound to each YdiB molecule. The final model for the asymmetric unit refined at 2.5 Å has an R_{work} of 22.6% and an R_{free} of 29.4% and consists of 4274 protein atoms, two NAD^+ cofactors, five phosphate ions, and 156 water molecules. The relatively high R_{free} value is likely explained by the presence of two overlapping conformations of helix α7 in molecule A, which render the electron density map difficult to model. In the vicinity of each molecule, several disordered electron density features were also left unassigned, because they do not respect the hydrogen bonding criteria of water molecules. The final structure was a good stereochemistry as evaluated using the PROCHECK program (26). The structure is deposited with the Protein Data Bank with the code 1O9B.

Biochemical Characterization—To evaluate the oligomeric state of YdiB, dynamic light-scattering (DLS) measurements were done on a

TABLE I
Crystallographic data, phasing, and refinement statistics

Data and phasing statistics					
Data set	AroE (native)	AroE $HgCN_2$	YdiB selenium peak	YdiB selenium edge	YdiB selenium remote
Space group	C2		$P6_4$		
Unit cell	a = 110.52Å b = 140.02Å c = 102.71Å β = 122.07°	a = 110.80Å b = 138.75Å c = 103.04Å β = 122.79°	$a = b$ = 157.874 Å, c = 40.011 Å		
Resolution (Å)	1.5	2.0	2.5		
Wavelength (Å)	0.87	1.0072	0.97913	0.97929	0.97161
Unique reflections	208, 460	173,047	37, 642[a]	37,647[a]	38,407[a]
Completeness (%)[b]	98.9 (86.7)	99.9 (100.0)	99.8 (100)	99.9 (100)	99.7 (100)
R_{merge} (AroE)/R_{sym} (YdiB) (%)[b,c]	4.8 (76.6)	2.9 (10.0)	3.5 (22.7)	3.7 (24.4)	2.8 (16.0)
Multiplicity	5.3	4.1	3.86	3.87	3.88
Mean $\langle I/\epsilon(I)\rangle 1$[b]	30.0 (1.5)	39.8 (14.2)	20.7 (5.1)	20.0 (4.9)	25.5 (7.2)
Phasing power acentric/centric		1.3/1.0	2.63 (2.45)	3.20 (2.89)	3.13 (2.66)
Mean figure of merit		0.46		0.67	
Refinement statistics					
Resolution range (Å)	87.7-1.5		20.0-2.5		
R_{work}/R_{free}[d]	14.7 (17.6)		22.6 (29.4)		
Unique reflections	208,460		20,136[e]		
No. of protein/other atoms	8256/1273(H_2O)/270 ($NADP^+$ + sulfate + DTT)		4274/156(H_2O)/112 (NAD^+ + phosphate)		
R.m.s.d. Å bonds (Å)	0.018		0.010		
R.m.s.d.[f] angles (°)	1.8		1.5		
Ramachandran plot					
Residues in most favored and additional allowed regions (%)	93.0/7.0		87.7/12.3		

[a] Unmerged and Bijvoet pairs.
[b] Number in parentheses represents values for highest resolution bin (1.539-1.500 Å AroE, 2.59-2.50 Å YdiB).
[c] $R_{merge} = \Sigma|I_{obs} - I_{avg}|/\Sigma I_{avg}$, $R_{sym} = \Sigma|I_{obs} - I_{avg}|/\Sigma I_{avg}$, where the summation is over all symmetry equivalent reflections.
[d] R_{free} was calculated with 9.9% (AroE) and 5.0% (YdiB) of reflections.
[e] Merged Bijvoet pairs.
[f] Root mean square deviation.

solution of YdiB concentrated at 1 and 11 mg/ml, in the presence or in the absence of 2 mM NADH. The measurements were performed using a DynaPro 801 instrument (Protein Solutions, Charlottesville, VA). To confirm these DLS results, gel filtration analysis was also performed using a Superdex 75 column (Amersham Biosciences) and calibrated with the reference protein mixture recommended by Amersham Biosciences. The YdiB sample (200 μl, 11 mg/ml) was injected and eluted at 1 ml/min in the same buffer (20 mM Tris-HCl, pH 7.5, 200 M NaCl, 5% (w/v) glycerol, 5 mM DTT). The enzymatic activities of AroE and YdiB were assayed at 20 °C by monitoring the reduction of NAD^+ or $NADP^+$ at 340 nm ($\epsilon = 6.18 \times 10^{-3}$ M^{-1} cm^{-1}) in the presence of either shikimic acid or quinic acid. To test possible inhibition by NAD^+, the enzymatic activity of AroE was assayed in the following buffers: 100 mM Tris-HCl, pH 9.0, 5 mM shikimic, 200 μM $NADP^+$, and 20 mM NAD^+. To measure the kinetic parameters for each cofactor, the assay mixture (total volume 200 μl) consisted of 100 mM Tris-HCl, pH 9.0, 5 mM shikimic or quinic acid, and six different values for the cofactor NAD^+ or $NADP^+$ (4, 2, and 1 mM and 500, 250, and 125 μM). Similarly to measure the kinetic parameters for both substrates, the assay mixture consisted of 100 mM Tris-HCl, pH 9.0, 5 mM NAD^+ or $NADP^+$, and six different values for shikimic or quinic acid (4, 2, and 1 mM and 500, 250, and 125 μM). To measure the activity, 10 μl of enzyme ($[AroE]_{stock}$ = 0.17 nM, $[YdiB]_{stock}$ = 800 nM) was added to the assay mixture. These enzyme concentrations were chosen in order to follow the initial reaction rate. The absorbance at 340 nm was measured for 30 min against a blank consisting of the assay mixture without enzyme. Each measure was taken in triplicate and simultaneously using a 96-well quartz plate. The kinetic parameters were deduced by the Lineweaver-Burk method. These reactions were monitored using the Plate Reader Spectra Max (Molecular Devices, Sunnyvale, CA). All chemicals were purchased from Sigma.

RESULTS AND DISCUSSION

Characterization of YdiB as a Dual Specificity Quinate/Shikimate Dehydrogenase—E. coli YdiB shows sequence similarity with AroE from the same organism (13) and with the quinate/shikimate dehydrogenases Qa-3 from *N. crassa* (29) and QutB from *Emericella nidulans* (30). We used this knowledge as a starting point to investigate its substrate specificity by monitoring the reduction of NAD^+ or $NADP^+$ in the presence of either shikimic acid or quinic acid. As a control, purified recombinant AroE protein was tested under the same conditions. As expected, AroE oxidized shikimic acid using $NADP^+$ as cofactor but displayed no activity in the presence of NAD^+. The kinetic parameters are very similar for both the cofactor and the substrate (NADP K_m = 56 μM, k_{cat} = 14,200 min^{-1}, shikimate K_m = 65 μM, and k_{cat} = 14,200 min^{-1}). Quinic acid, even at a high concentration of 5 mM, is not a substrate for AroE, either in the presence of $NADP^+$ or NAD^+. To determine whether NAD^+ acts as a competitive inhibitor with respect to $NADP^+$, the oxidation of shikimic acid was assayed with a $NADP^+/NAD^+$ ratio of 1:100 (200 μM $NADP^+$ and 20 mM NAD^+). Despite the large excess of NAD^+, the activity of AroE was not significantly altered, indicating that NAD^+ is not a competitive inhibitor of $NADP^+$ and that AroE likely does not bind NAD^+.

In contrast, YdiB is able to oxidize shikimic acid by using either $NADP^+$ or NAD^+ as cofactor. At saturation of shikimate, YdiB displays similar kinetic parameters for both cofactors ($NADP^+$, K_m = 100 μM, k_{cat} = 7 min^{-1}; NAD^+, K_m = 87 μM, k_{cat} = 3 min^{-1}). The K_m values significantly differ for the shikimic acid, according to the type of cofactor used at saturation: shikimate + $NADP^+$, K_m = 120 μM, k_{cat} = 7 min^{-1}; shikimate + NAD^+ K_m = 20 μM, k_{cat} = 3 min^{-1}. Contrary to AroE, YdiB also displays a clear activity on quinic acid, with either $NADP^+$ or NAD^+ as a cofactor. At saturation of quinate, YdiB displays a five times lower K_m for NAD^+ (K_m = 116 μM, k_{cat} = 3 min^{-1}) than for $NADP^+$ (K_m = 500 μM, k_{cat} = 3 min^{-1}). This phenomenon is accentuated for the K_m of quinic acid, which is 10 times lower at saturation of NAD^+ (K_m = 41 μM, k_{cat} = 3 min^{-1}) than at saturation of $NADP^+$ (K_m = 555 μM, k_{cat} = 3 min^{-1}).

YdiB is therefore the first quinate/shikimate dehydrogenase identified in *E. coli*. Although this enzyme has a lower catalytic efficiency (~2–4000-fold) compared with that of AroE, this is compensated by a broader substrate and cofactor specificity. The low specific activity of YdiB likely explains why it was not identified alongside AroE during the initial purification of this activity from *E. coli* (12). Although it is clear that YdiB is NADP/NAD-dependent dehydrogenase, we cannot exclude the possibility that its physiological substrate is neither shikimate nor quinate, considering its low catalytic efficiency.

Nevertheless, AroE and YdiB display a fairly equivalent affinity for their ligands, as shown by the similar range of their K_m values. Furthermore, YdiB seems equally active on shikimic and quinic acid, because their K_m values are comparable in the presence of NAD^+ (20 and 40 μM, respectively). In contrast, the behavior of YdiB is different according to which cofactor is used. YdiB has a tendency to be more "efficient" in the presence of NAD^+, as shown by the discrepancy between the K_m values for shikimate/quinate at the saturation of either NAD^+ (20/40 μM) or $NADP^+$ (120/555 μM). This difference could be explained by a lower affinity for $NADP^+$, as shown by the cofactor K_m in the presence of quinate or by a binding of $NADP^+$ in a less productive manner (shikimate case).

Overall Structure of E. coli AroE and YdiB—The asymmetric unit of AroE crystals contains four protein molecules (Met^1–Ser^{271}) complexed with $NADP^+$, and a total of 13 sulfate ions, 1277 water molecules, and 1 molecule of DTT bound in the active site of molecule A. The four protein molecules are related by pseudo 222 symmetry as reported previously (19). The YdiB asymmetric unit comprises two molecules (Tyr^7–Phe^{286}), related by 2-fold noncrystallographic symmetry, each complexed with NAD^+, 2 phosphate ions, and 156 water molecules. The residues Met^1–Lys^6 and Gly^{287}–Ala^{288} are disordered and are not included in the model. Unless specified otherwise, we will refer to residues according to AroE numbering, with those of YdiB referenced in parentheses.

Despite the relatively low sequence similarity between AroE and YdiB, the two enzymes have highly similar structures that adopt the same fold (Figs. 2 and 3). The molecules have a somewhat elongated shape (55 × 40 × 30 Å) and comprise two domains. The first domain is made of two discontinuous segments, Met^1–Thr^{101} (Met^7–Thr^{106}) and Gly^{237}–Ser^{271} (Gly^{255}–Phe^{286}), whereas the second domain encompasses Gly^{119}–Asp^{236} (Gly^{124}–Asp^{254}). Both domains have α/β architectures and are connected by the helix α5 and a short linker, Asp^{102}–Pro^{118} (Asp^{107}–Lys^{123}). The arrangement of these two domains along the connecting helices creates a deep groove in which the cofactor $NADP^+$ (or NAD^+) is located (Fig. 2).

The N-terminal domain consists of a mainly parallel six-stranded β-sheet and six α-helices. The strand order is 2-1-3-5-6-4, with the strand β5 being antiparallel to the other strands. The first three β-strands follow a regular β/α succession, with the helices α1 and α2 parallel to the β-strands, flanking opposite sides of the sheet. The next $\alpha/\beta/\alpha$ unit is irregular, with the helix α3 oriented at ~45° relative to the direction of the sheet, and the short, one turn helix α4 nearly perpendicular to the strand β4. The domain is completed by a C-terminal α-helical hairpin (α9 and α10), which packs against the β-sheet on the same side as α1. According to the DALI algorithm (31), this domain shows topological and structural similarity with the C-terminal domain of glycyl-tRNA synthetase (Protein Data Bank code 1ATI), which has strand order 2-1-3-4-5 (β4 antiparallel to the other strands). Out of 129 residues, 80 Cα atoms can be superimposed on AroE with r.m.s.d. of 2.5 Å. In AroE/YdiB the extended loop between strands β3 and β5 contains two helices and folds back onto the

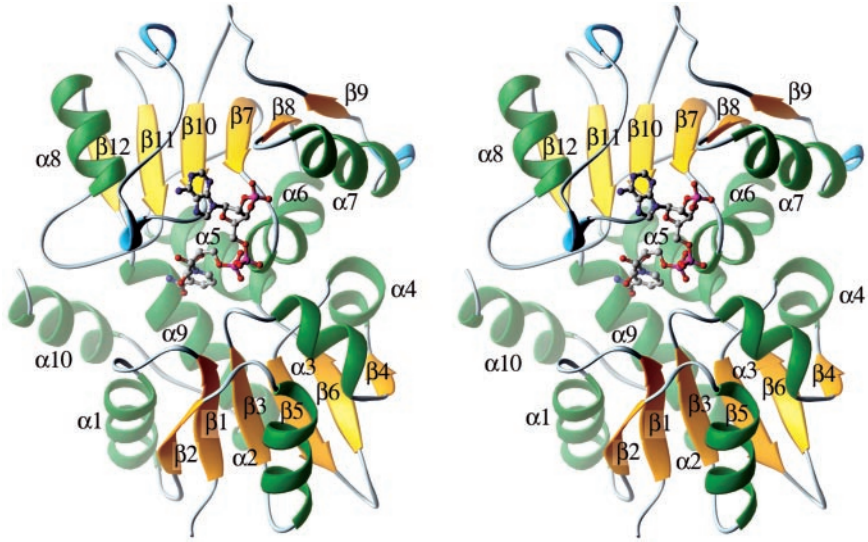

FIG. 2. **A stereo ribbon representation of *E. coli* AroE (chain A) shows the overall fold of this enzyme with β strands colored in *gold*, α-helices in *green*, and 3_{10} helices in *turquoise*.** The cofactor $NADP^+$ is shown in a *ball* and *stick* representation, with nitrogen colored *blue*, oxygen colored *red*, carbon colored *gray*, and phosphorus colored *magenta*. Figs. 2 and 6 were prepared using the program RIBBONS (47).

β-sheet adding a sixth strand (β4) at the end. The corresponding loop in glycyl-tRNA synthetase is several residues shorter and extends away from the β-sheet. The fold of AroE/YdiB is also similar to the N-terminal part of the molybdenum cofactor biosynthesis protein MogA (Protein Data Bank code 1DI6, r.m.s.d. of 3.1 Å over 102 residues). Although the two folds differ in the strand order (2-1-3-6-5-4 in MogA with β5 antiparallel to the other strands), corresponding to a switch in the relative positions of strands β5 and β6, there is additionally a good spatial overlap of several helices.

The C-terminal domain or NAD(P)-binding domain could not be recognized from its amino acid sequence; however, this domain adopts a nearly canonical Rossmann fold, *i.e.* a six-stranded parallel β-sheet, with the strand order 3-2-1-4-5-6, and α-helices on both sides parallel to the β-strands. The fourth α-helix present in the canonical Rossmann fold is missing in YdiB, whereas the third and fourth α-helices are replaced by irregular loops in AroE. As a result, the AroE/YdiB NAD(P)-binding domains are among some of the shortest reported, sharing most structural homology with *S*-adenosylhomocysteine hydrolase (Protein Data Bank code 1D4G, r.m.s.d. 1.64 Å over 153 Cα atoms) and mouse class II alcohol dehydrogenase (Protein Data Bank code 1E3L, r.m.s.d. 1.87 Å over 160 Cα atoms). The SDH family provides a new example of a protein family displaying the dinucleotide binding fold, without significant sequence homology with other Rossmann fold families; this may indicate early divergence from the ancestral fold.

Quaternary Structures of AroE and YdiB—Whereas AroE has been shown to be a monomeric protein (12, 19), dynamic light scattering measurements on YdiB using different protein concentrations (1 and 11 mg/ml), both in the presence and absence of NADH, show that YdiB has a hydrodynamic radius consistent with a particle of ~60 kDa, indicating that this protein forms dimers. This was verified by size exclusion chromatography where the apoprotein eluted as a single species of 64 kDa. Analysis of the different protein-protein interfaces within the crystal structure of YdiB shows that the largest contact surface area is between the two molecules in the asymmetric unit. The two monomers are related by pseudo 2-fold symmetry, with the dimer interface formed by residues from strands β1, β2, and the helix α2 of the two N-terminal domains. This head-to-head packing of the N-terminal domains creates a highly elongated dimer with diametrically positioned active site clefts. The interface involves contacts made by 16 residues from each molecule and is predominantly hydrophobic in nature. The dimer buries 1400 $Å^2$ of solvent-accessible surface area (700 $Å^2$ from each monomer), which is at the lower end of values observed for protein-protein interfaces (32). Such an interface is not without precedent as a much smaller, solely hydrophobic, interface has been observed for the structure of Ocr from bacteriophage T7 (33). If the YdiB dimer interface has been correctly identified, then the hydrophobic residues forming this interface are mostly replaced by polar or smaller amino acids in AroE, notably YdiB (AroE) Leu^9 (Thr^3), Met^{40} (Gly^{34}), Phe^{42} (Val^{36}), Leu^{59} (Ala^{53}), and Met^{61} (Gly^{55}) (Fig. 3). These amino acid substitutions eliminate the hydrophobic patch on the surface of the YdiB monomer, giving a more hydrophilic character to the N-terminal domain of AroE and explaining why this protein is present as a monomer in solution.

The Cofactor Binding Site of AroE and YdiB—In all AroE and YdiB molecules the electron density for the cofactor, NAD^+ in YdiB and $NADP^+$ for AroE, is very well defined (Fig. 4). In the following description we will refer to molecule B of YdiB and molecule A of AroE as these have the lowest average *B*-factors. The $NAD(P)^+$ cofactor is located outside the carboxyl ends of β-strands β7–β10 at a switch point in the central β-sheet of the C-terminal domain. The superposition of the C-terminal domains of AroE and YdiB results in good superposition of NAD^+ and $NADP^+$, especially of their diphosphate groups and nicotinamide rings. A somewhat larger difference, an ~2-Å shift, occurs in the relative position of the adenosine.

Similar Recognition of Nicotinamide and Pyrophosphate—The binding of the nicotinamide and pyrophosphate moieties is similar in AroE and YdiB. The amide group N-7 of the nicotinamide ring is hydrogen-bonded to the carbonyl group of two residues, Met^{213} (Cys^{232}) and the invariant Gly^{237} (Gly^{255}) (Fig. 3). The neighboring ribose forms only van der Waals contacts to the hydrophobic side chains. The pyrophosphate moiety contacts the glycine-rich loop that connects strand β7 and helix α6 (Fig. 4) and forms hydrogen bonds to the backbone N atoms of Gly^{129} and Ala^{130} (Gly^{134} and Ala^{135}). A sequence pattern G [A,s,g] G G [A,t] [A,S,g] corresponding to the diphosphate-binding loop is conserved in the entire SDH family (Fig. 3). This fingerprint is yet another modification of the canonical pattern identified in NAD-dependent dehydrogenases: G-S2-S3-G-S5-S6-G, where S2 may be absent, S3 and S5 are variable, and S6 is always a hydrophobic residue, whose side chain is directed toward the nicotinamide moiety (34). With the missing residue S2, the main differences in the AroE fingerprint are the strict conservation of a glycine at the usually variable position S5, the presence of a less hydrophobic residue at position S6, and a small residue in place of Gly at the next position.

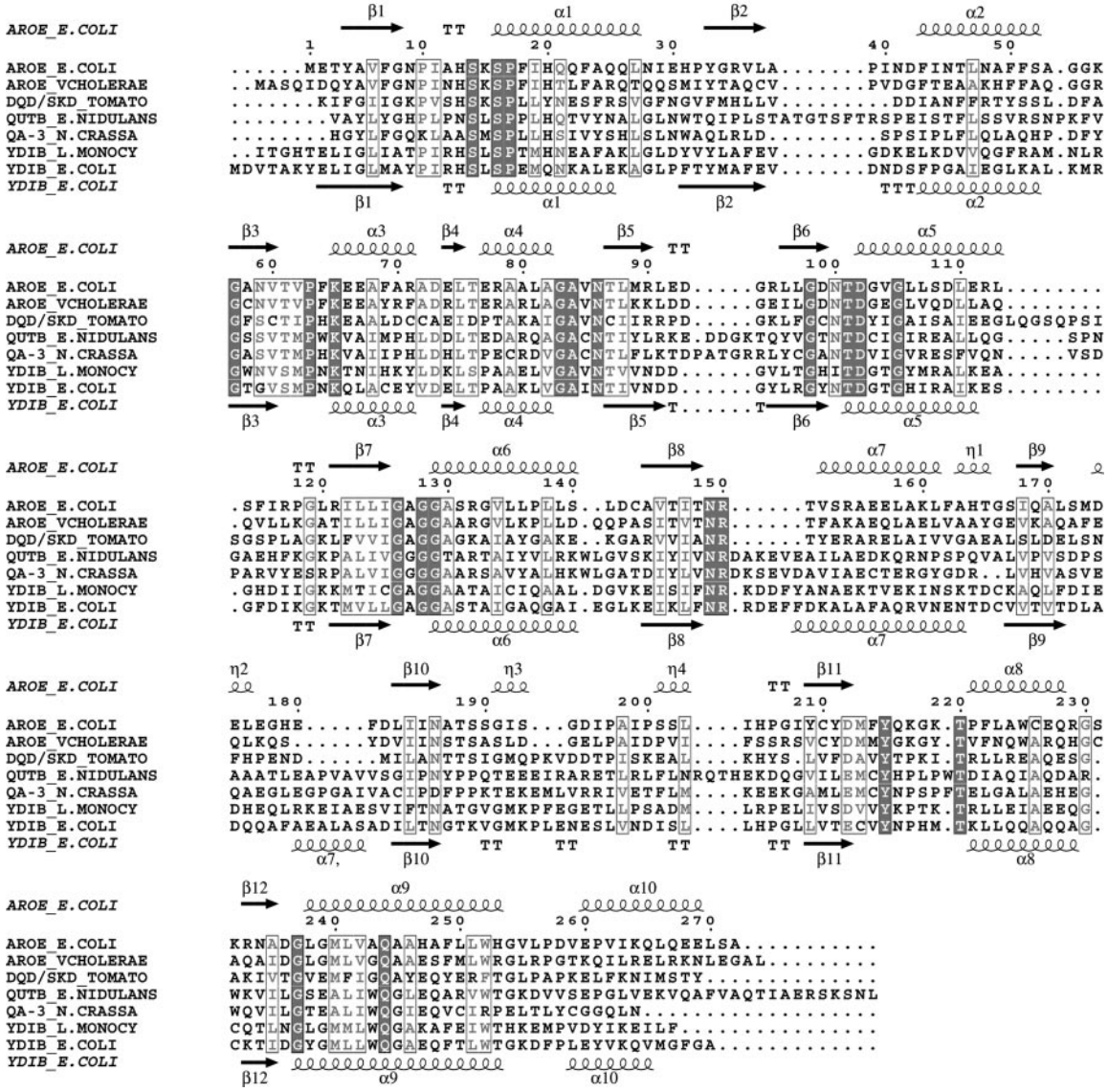

FIG. 3. **Structure-based sequence alignment of the SDH family.** Representative shikimate dehydrogenase and quinate dehydrogenase sequences from bacteria, yeast, and plants are compared with the secondary structures of *E. coli* AroE and YdiB. α-Helices and β-strands are represented as helices and *arrows,* respectively, and β-turns are marked with *TT*. This sequence alignment was created using the following sequences from GenBank™: *E. coli* K12 AroE (NP_417740, residues 1–272), *Vibrio cholerae* AroE (Q9KVT3, 1–278), *Lycopersicon esculentum* (tomato) dehydroquinate dehydratase/shikimate dehydrogenase (T06264, 271–545), *E. nidulans* QutB (CAA31880, 13–329), *N. crassa* Qa-3 (CAA32751, 40–321), *L. monocytogenes* YdiB (Q8Y9N5, 5–291), and *E. coli* K12 YdiB (NP_416207, 1–288). *Dark shaded boxes* enclose conserved positions, and *light shaded boxes* show positions with homologous residues. Figure was prepared using the program ESPript (48).

Cofactor Specificity Determinants within the Adenosine-binding Pocket—In contrast to the vast majority of the NAD(P)-dependent dehydrogenases, which have a strong specificity for either NAD or NADP (34), members of the SDH family show a diversity of cofactor specificity. *E. coli* AroE, involved in biosynthesis, is strictly NADP-dependent (12), whereas *N. crassa* Qa-3 and *E. nidulans* QutB display a strong preference for NAD (29, 30), and *E. coli* YdiB is able to use both cofactors. Therefore, the comparison of the cofactor binding sites in AroE and YdiB is of interest as it reveals the structural features necessary to discriminate between NADP and NAD in the SDH family.

The binding of the adenine moiety by both enzymes is typical for NADP-dependent dehydrogenases as it contains an arginine side chain that stacks against the adenine ring and lacks a carboxylic residue (replaced by Asn) that chelates the diol group of the ribose in NAD complexes (34). In the SDH family the loop between strand β8 and helix α7 features two strictly conserved residues, Asn^{149} and Arg^{150} (Asn^{155} and Arg^{156}, Fig. 3), which are both involved in the recognition of the adenosine moiety. There are, however, differences in their interactions with the cofactor in the two enzymes. In the AroE-$NADP^+$ complex, the hydroxyl group O-3′ of the adenosine ribose is hydrogen-bonded to $Asn^{149(OD1)}$, as well as to the main chain NH of Ala^{127}, located in the glycine-rich loop (Fig. 4*A*). In addition, the amide of Asn^{149} forms a hydrogen bond to the O-1 atom of the 2′-phosphate. Arg^{150} forms two hydrogen bonds with the other oxygen atoms of the phosphate substituent, whereas its guanidinium group stacks against the A-face of the adenine ring. This phosphate is further stabilized by electrostatic interactions with Arg^{154} from helix α7 and by a hydrogen bond with $Thr^{151(OH)}$. Face B of adenine contacts the side chain of Thr^{188} and Ser^{190} (Fig. 4*A*). The arginines 150 and 154 play a crucial role in adenosine phosphate binding as they form an "electrostatic clamp" that sandwiches the phosphate substituent.

In YdiB there are several substitutions affecting the interactions with NAD^+ (Fig. 4*B*). Val^{206}, which replaces Ser^{190} of

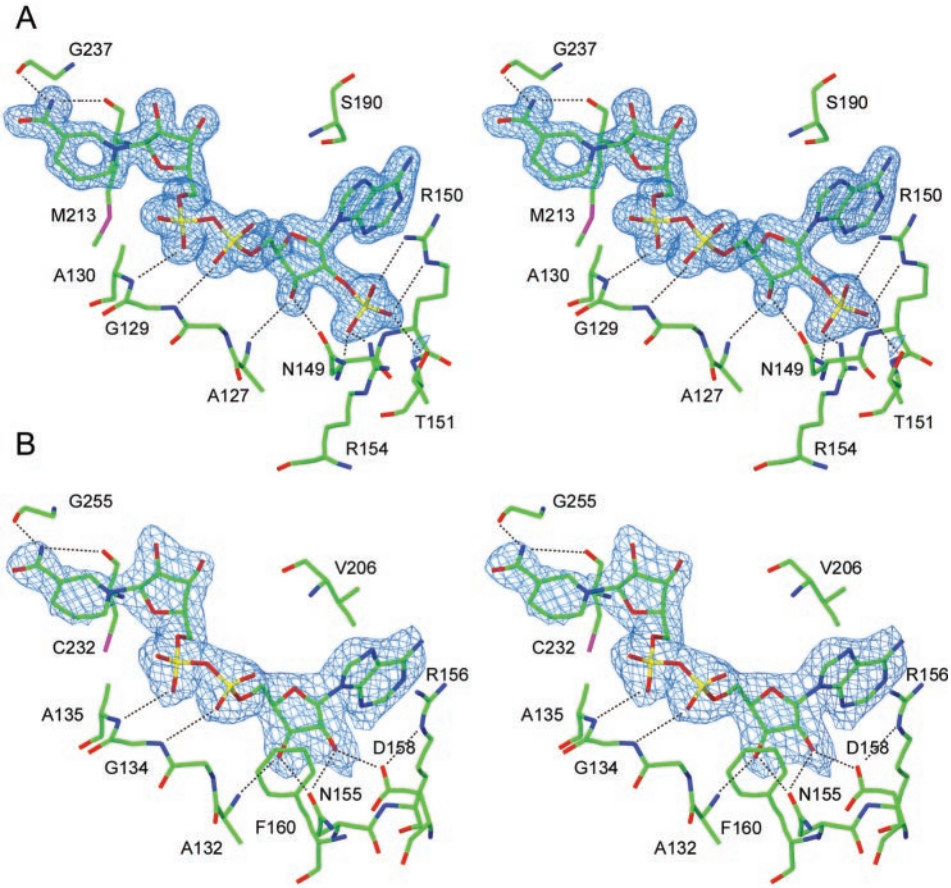

FIG. 4. **Comparison of the cofactor-binding site of AroE (*A*) and YdiB (*B*).** Stereo view of the $2F_o - F_c$ OMIT maps based on the final refined coordinates of AroE (1.5 Å resolution, 1.5 σ) and YdiB (2.5 Å resolution, 1.0 σ), excluding the atoms in a radius of 3.5 Å around $NADP^+$ and NAD^+, respectively. The cofactors and the important amino acids are represented in *stick*, with carbons colored *green*, oxygen colored *red*, nitrogen colored *blue*, sulfur colored *magenta,* and phosphate colored *yellow*. The hydrogen bonds are shown as *dashed lines*. Figure was prepared using the program O (28).

AroE, orients its aliphatic side chain perpendicularly to the B-face of the adenine ring, forming a CH-π-electron hydrogen bond (35). The bulging of Val^{206} is accompanied by a compensating shift of Arg^{156}, which maintains its stacking against the A-face of a slightly translated adenine. This arrangement provides for hydrogen bonds of O-2′ and O-3′ of NAD^+ ribose to $Asn^{155(OD1)}$ as well as the O-3′ to a backbone NH of Ala^{132} (Fig. 4*B*). The NAD^+ binding in YdiB is favored by the substitution of Thr^{151} and Arg^{154} of AroE by Asp^{158} and Phe^{160}, respectively. Asp^{158} is hydrogen-bonded to the hydroxyl group O-2′ of the ribose and also stabilizes Arg^{156} through a salt bridge. The hydrophobic residue Phe^{160} creates a neutral environment, which is less discriminating than the basic binding pocket observed in the AroE structure (Fig. 5, *B* and *C*). The capacity of YdiB to also bind $NADP^+$ likely involves a conformational change of Asp^{158} to avoid electrostatic repulsion with the phosphate group. A low resolution structure of YdiB co-crystallized with NADP confirmed that this cofactor is located in a position similar to that of NAD^+. The loop β8-α7, which contains Asp^{158}, is displaced in this structure in order to provide a phosphate-binding site and as a result is poorly ordered.

The Active Site and Its Conformational Flexibility—The substrate-binding site is identified by the position of the nicotinamide ring of the cofactor and is delineated almost entirely by residues from the N-terminal domain. The binding site is in a pocket formed by the C-terminal ends of the β-strands, the N-terminal end of helix α1, the side of helix α9, the extended loop between β1 and α1, and the first residues from the connecting helix α5. Most of the residues absolutely conserved in the SDH family are located in this pocket, *i.e.* Ser^{14}, Ser^{16}, Lys^{65}, Asn^{86}, Thr^{101}, Asp^{102}, and Gln^{244}. At position 61 (67), a serine or a threonine is also always observed in the SDH family (Fig. 3). A sulfate or phosphate ion is present in this cavity in all AroE and YdiB molecules. In molecules A and B of AroE, this anion is located at the top of the pocket and is hydrogen-bonded to the hydroxyl groups of Ser^{14}, Ser^{16}, Thr^{61}, and Tyr^{215} (Fig. 6*A*), whereas in the remaining AroE and YdiB molecules it lies at the bottom of the cavity, hydrogen-bonded to the side chains of Lys^{65} and Thr^{61} (Lys^{71} and Ser^{67}). In molecule A of AroE, a DTT molecule is also present in this pocket, tightly bound through numerous hydrogen bonds involving its thiol and hydroxyl groups to the conserved AroE residues: DTT^{SH1}–$Gln^{244(OE1)}$, DTT^{OH2}–$Lys^{69(NZ)}$, $Asn^{86(ND2)}$ and $Asp^{102O(D1)}$, and DTT^{SH4}–$Thr^{61O(G)}$ (Fig. 6*A*).

The comparison of independent molecules of AroE and YdiB shows clear differences in the relative disposition of their domains. Three different conformations are observed for AroE (molecules A/B, C, and D), whereas the two molecules of YdiB display similar conformation. Comparing individual domains of the same protein gives an r.m.s.d. in the range of 0.3–0.6 Å. Superposition of the individual domains of AroE and YdiB results in an r.m.s.d. of ~1.3 Å for 104 of 136 Cα atoms and ~1.4 Å for 100 of 137 Cα atoms, for the substrate- and cofactor-binding domains, respectively. However, these numbers for the entire molecules are significantly larger, 1.2–1.6 Å for the independent AroE molecules and 2.8–3.3 Å for the comparison of AroE and YdiB molecules (Fig. 5*A*). Among these conformations, molecule A of AroE represent the most "closed" form,

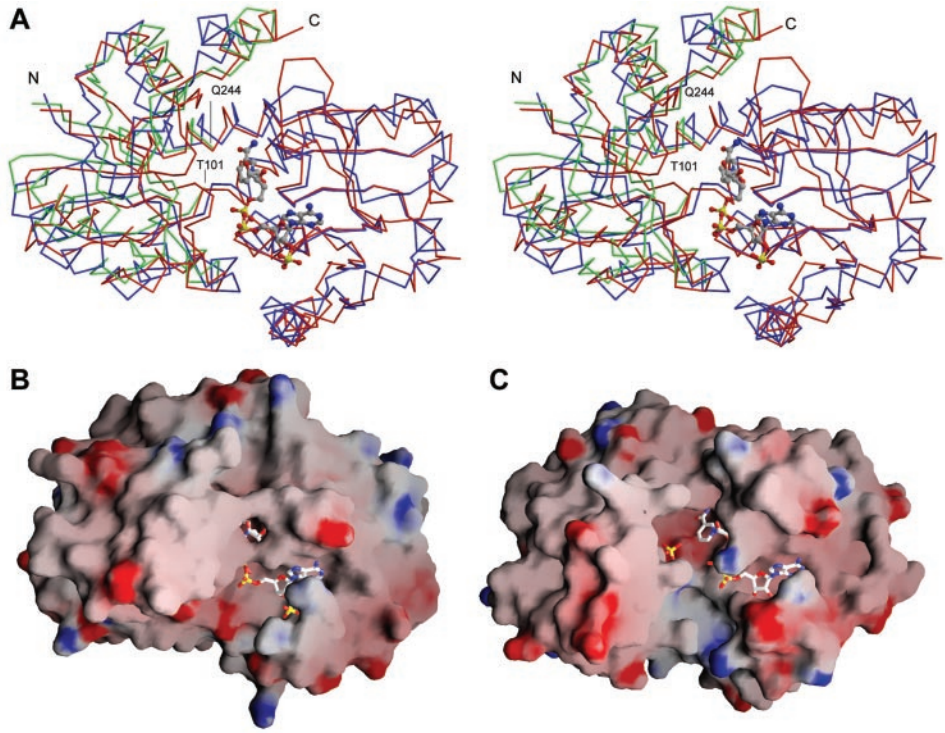

FIG. 5. *A*, stereo view of the Cα trace of independent molecules of AroE and YdiB superimposed on the nucleotide-binding domain of the enzyme with the NADPH shown in *ball* and *stick*. The C-terminal NAD(P)-binding domains are only shown for molecules A of AroE (*red*), which represents a closed conformation of the shikimate dehydrogenase, and of YdiB (*blue*), which represent an open conformation. The molecule C of AroE (*green*) represents an intermediary conformation. The two residues that act as hinges for the conformational change are indicated (AroE numbering). Surface representations of the molecules A of AroE, in closed conformation (*B*), and of YdiB, in open conformation (*C*), are colored according to electrostatic potential with $NADP^+$ and NAD^+ shown in *ball* and *stick* representation. Figure was prepared with the programs MOLSCRIPT (49) and GRASP (50).

whereas molecule A of YdiB represent the most "open" form of the enzyme (Fig. 5, *B* and *C*). The transition between these two extreme conformations corresponds to a rotation of ~25° around an axis passing approximately through the Cα of Gln^{26} (Lys^{32}) and Asp^{102} (Asp^{107}). Consequently, the tip of the N-terminal domain traverses a distance of ~14 Å between the open and closed structures.

This overall conformational change is concomitant with the rearrangement of the hydrogen bonding network in the junction region between the N- and C-terminal domains. Among the five residues involved in this network, three (Asn^{86}, Thr^{101}, and Gln^{244} (Asn^{92}, Thr^{106}, and Gln^{262})) are invariant in the SDH family, whereas a fourth residue, Thr^{87} (Thr^{93}), is conserved in 92% of sequences. The last residue, Asn^{59}, is conserved in ~60% of the sequences and is substituted by small residues (Gly, Ala, and Ser) in the remainder of the SDH family. In this latter group, which includes YdiB, we find a compensating replacement of Ala^{248} by a glutamine (Gln^{266}), whose carboxyamide group overlaps that of Asn^{59}, resulting in a spatial invariance of a polar group at this position. In the open conformation, all these residues are linked by hydrogen bonding interactions between their side chains: $Asn^{59(NE1)}$ (Gln^{266})–$Thr^{87O(G)}$–$Asn^{86O(D1)}$, $Asn^{86(ND2)}$–$Thr^{101O(G)}$–$Gln^{244(NE2)}$-$Asn^{59O(E1)}$ (Gln^{266}). This circular network rearranges in the closed conformation, as Gln^{244} is no longer hydrogen-bonded to the side chains of Asn^{59} (Gln^{266}) and Thr^{101}. Instead, this glutamine side chain makes a hydrogen bond to the side chain of Asn^{86}, whereas its main chain carbonyl group is hydrogen-bonded to $Thr^{101(OH)}$. Because the closed conformation was found in the molecule that binds DTT, we speculate that the conformational change, which closes the central cleft, occurs upon substrate binding and is necessary for the formation of a productive active site. The cluster of conserved residues in the junction region therefore acts as a hinge, stabilizing the open conformation at the beginning of a catalytic cycle and then favoring the closing of the active site cleft when the substrate is present.

The Reaction Mechanism of Shikimate/Quinate Dehydrogenase—The presence in the closed active site of a DTT molecule and a sulfate ion, contacting invariant residues, suggests the possible interactions between shikimate dehydrogenase and its substrate (Fig. 6*A*). The integration of the sequence and biochemical and structural evidence led us to propose a model for the recognition of 3-dehydroshikimate (Fig. 6*B*). The enzyme catalyzes the stereospecific reduction of 3-dehydroshikimate to shikimate and, as such, requires precise positioning of the substrate. It was shown that hydride transfer occurs from the A-side of NADPH (36), which is consistent with the orientation of the cofactor in the active sites of the two structures. For catalysis to occur, the C-3 of 3-dehydroshikimate/3-dehydroquinate must be positioned to receive the hydrogen from C-4 of the nicotinamide ring. The location of the C-3 and C-4 of DTT in the vicinity of the C-4 of $NADP^+$ is consistent with such positioning (Fig. 6*A*).

At the same time, we expect that the carboxylate would form specific interactions within the substrate-binding pocket. In the other enzymes in the shikimate pathway, the carboxylate of the substrate is bound by either an arginine (type I dehydroquinase (8), dehydroquinate synthase (7), 5-enolpyruvylshikimate-3-phosphate synthase (11)) or main chain amides (type II dehydroquinase (37)). Given the position of the conserved residues within the active site, the loop between β1 and α1 delineated by two strongly conserved proline residues adopts a conformation capable of binding a carboxylate in a similar manner to the type II dehydroquinase. The conserved serine residues at positions 14 and 16 most likely contribute to carboxylate bind-

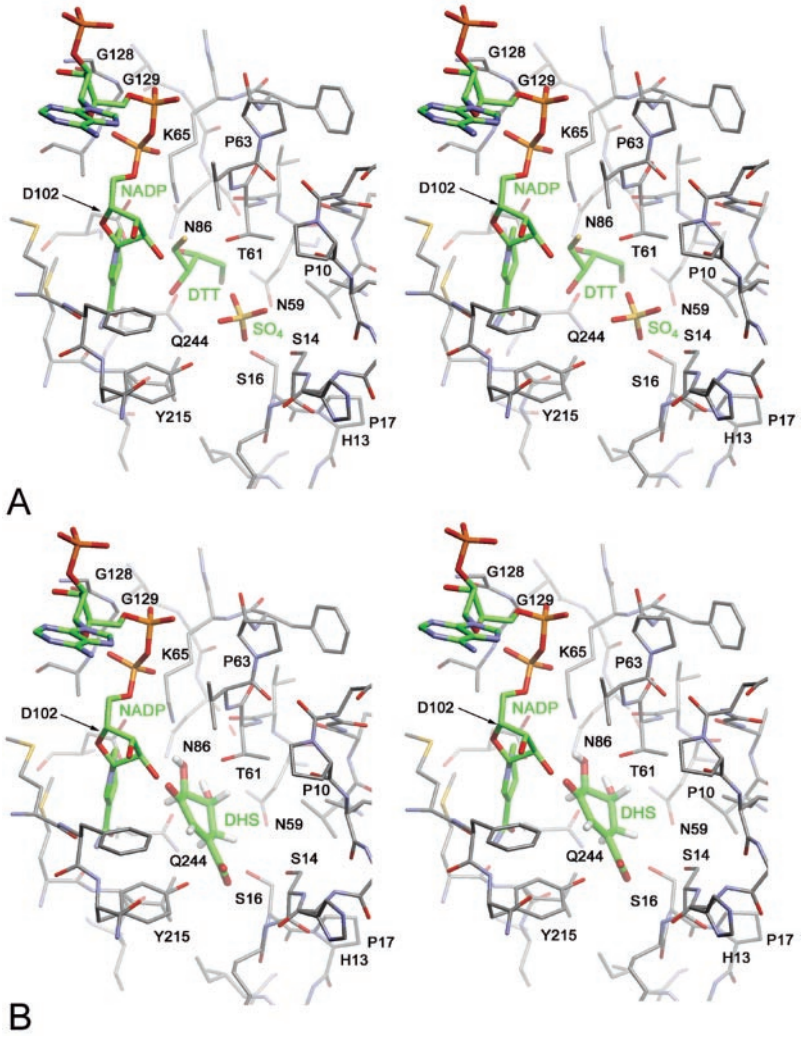

FIG. 6. *A*, stereo view of AroE active site with $NADP^+$, bound sulfate, and DTT molecules represented in *thick stick* (carbon atoms colored *green*) and protein atoms represented in *thinner stick* (carbon atoms colored *gray*). All atoms are colored according to atom type, and conserved amino acid residues are labeled. *B*, a molecular model of the binding of dehydroshikimate to the active site of AroE. This model does not represent the proposed ternary complex as further closing of the active site is envisaged and cannot be modeled with confidence. The model serves instead as a guide to the chemically correct orientation of the substrate which is necessary for catalysis to occur.

ing so that both carboxyl oxygens form two hydrogen bonds to the protein. The conserved tyrosine 215, located in the C-terminal domain but whose side chain points toward the substrate pocket, is also likely to establish an additional hydrogen bond with the carboxylate. The location of the sulfate ion in this region of the structure and its interactions with these three conserved residues support this hypothesis (Fig. 6*A*).

The substrate in this orientation will form hydrogen bonds between C-4 hydroxyl and the side chain of Lys^{65} and Asp^{102}, whereas the C-5 hydroxyl would be positioned down into the active site forming hydrogen bonds with Gln^{244}. Such hydrogen bonds are observed between AroE and the groups OH2 and SH1 of DTT (Fig. 6*A*). Previous studies of *Pisum sativum* shikimate dehydrogenase have shown that substrate-like inhibitors of the enzyme require a C-4 hydroxyl, whereas either a C-5 hydroxyl or carboxylate group is needed for strong binding (38). By using a series of analogs of 3-dehydroshikimate that lack the C-4 and C-5 hydroxyls, Bugg and co-workers (39) demonstrated that the C-5 hydroxyl of the substrate has little effect on the specificity of *E. coli* AroE, whereas the C-4 hydroxyl is very significant. An estimation of the binding energy (based on k_{cat}/K_m (40)) between the C-4 hydroxyl and the enzyme suggests that this hydroxyl forms a hydrogen bond to a charged group (39). From the pH/rate profile of AroE, it has been suggested that this charged group is either a cysteine or an α-amino group (41). In this light, Lys^{65} (Lys^{71}) seems a good candidate for the residue coordinating the C-4 hydroxyl group.

By analogy with lactate dehydrogenase (42), an acid/base catalytic group is needed to donate a proton to the carbonyl of 3-dehydroshikimate during reduction and to remove a proton during oxidation of shikimate. The invariant Lys^{65} and Asp^{102} are the most likely candidates to assume this role, considering their proximity to both the nicotinamide ring and the SH4 and OH3 groups of DTT (Fig. 6*A*). Another possibility is the involvement of Thr^{61}, the 2′-hydroxyl of the cofactor, and His^{13} in a proton relay analogous to that found in alcohol dehydrogenase (43). The pH dependence of AroE (maximum at pH 7.3) is consistent with a histidine being involved in the mechanism. Histidine-specific chemical modification of AroE by diethylpyrocarbonate at pH 7.0 has been shown to inactivate the enzyme in a time-dependent manner, which was monitored by electrospray mass spectrometry (44). Two of the histidine residues in AroE could be protected from diethylpyrocarbonate modification by the presence of shikimate; one of these was identified as His^{13} (45). However, His^{13} is a significant distance from the 2′-hydroxyl of the cofactor (~11 Å), even in the closed conformation of AroE, and is not strictly conserved in the SDH family making it a less likely candidate for the catalytic acid/base.

The Evolutionary and Metabolic Implications from the Presence of Two Shikimate Dehydrogenase Genes in the E. coli Genome—The presence of two shikimate dehydrogenase isoforms in *E. coli* raises intriguing questions concerning their specific biological roles. The existence of a second shikimate dehydrogenase also affects the design of any potential drugs, because YdiB may compensate for the inhibition of AroE. Although the substrate specificity of YdiB has been identified

here, it is not yet clear if YdiB participates in the shikimate pathway or has another biological function. A systematic analysis of the bacterial genomes presently listed in the TIGR data base (www.tigr.org/) revealed that 14 species possess at least two shikimate dehydrogenase isozymes located at distinct loci. These microorganisms belong to distant phyla (*e.g.* α- and γ-Proteobacteria, Deinococcus-Thermus, Actinobacteria, and Firmicute), showing that this phenomenon is not limited only to *E. coli* or related species. Most of these homologous proteins display a similar, relatively low sequence identity to AroE and YdiB (20–30%), making a one-to-one assignment difficult. However, a few proteins are clearly either AroE-like (*Haemophilus influenzae* HI0655, *Pasteurella multicida* AroE, *Salmonella enterica* STY4396, *Salmonella typhimurium* STM3401, and *Yersinia pestis* YP00246) or YdiB-like (*Listeria innocua* Lin2338 and Lin0493, *Listeria monocytogenes* Lmo2236 and Lmo0490, *S. typhimurium* STM1359, and *Streptococcus pyogenes* SpyM18–1592) with sequence identity varying between 45 and 90%.

The location of the *aroE* and *ydiB* genes in the *E. coli* genome is also informative. A*roE* is flanked by genes of unknown or putative functions (*Yrd B, C,* and *D*), unrelated to the shikimate pathway. In contrast, *ydiB* is located between the gene *b1691*, coding a putative amino acid transport protein, and the gene *aroD*, coding type I 3-dehydroquinase. According to the Regulon DB data base (46), AroE and YdiB are independently regulated, whereas the cluster *b1691-ydiB-aroD* is under the control of the same promoter. Such organization of *ydiB* and *aroD* in one operon is also observed in pathogenic bacteria *L. innocua, L. monocytogenes* ($Gram^+$), and *S. typhimurium* ($Gram^-$). Moreover, these enzymes recognize the same substrate, 3-dehydroquinate. Therefore, YdiB may have a physiological role connected to that of *aroD*. More to the point, shikimate dehydrogenase and 3-dehydroquinase activities coassembled into a bifunctional protein in some plants and bacteria. Such a bifunctional enzyme could have evolved by the fusion of an ancestral bacterial *ydiB-aroD* gene cluster. Type I dehydroquinases like AroD are associated with biosynthesis, whereas type II dehydroquinases are known to function in synthetic and degradative pathways. The association of *ydiB-aroD* in one operon would therefore suggest its involvement in the shikimate pathway. In contrast, the substrate and cofactor promiscuity of YdiB would speak in favor of a different role. All presently known NAD-dependent quinate/shikimate dehydrogenases are involved in the catabolic quinate pathway. Therefore, YdiB may be essential for growth of *E. coli* with quinate as a sole carbon source (16), thus indicating the presence of a quinate pathway in this organism.

Acknowledgments—We thank J. Green for assistance in the purification of AroE and Prof. N. Price for critical reading of the manuscript. We are also grateful to R. Larocque for help with cloning the *ydiB* gene, Dr. S. Raymond for assistance in computational aspects, and Dr. J. D. Schrag for the comments on the manuscript.

REFERENCES

1. Herrmann, K. M., and Weaver, L. M. (1999) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50,** 473–503
2. Roberts, F., Roberts, C. W., Johnson, J. J., Kyle, D. E., Krell, T., Coggins, J. R., Coombs, G. H., Milhous, W. K., Tzipori, S., Ferguson, D. J., Chakrabarti, D., and McLeod, R. (1998) *Nature* **393,** 801–805
3. Kishore, G. M., and Shah, D. M. (1988) *Annu. Rev. Biochem.* **57,** 627–663
4. Davies, G. M., Barrett-Bee, K. J., Jude, D. A., Lehan, M., Nichols, W. W., Pinder, P. E., Thain, J. L., Watkins, W. J., and Wilson, R. G. (1994) *Antimicrob. Agents Chemother.* **38,** 403–406
5. Steinrücken, H. C., and Armhein, N. (1980) *Biochem. Biophys. Res. Commun.* **94,** 1207–1212
6. Baerson, S. R., Rodriguez, D. J., Tran, M., Feng, Y., Biest, N. A., and Dill, G. M. (2002) *Plant Physiol.* **129,** 1265–1275
7. Carpenter, E. P., Hawkins, A. R., Frost, J. W., and Brown, K. A. (1998) *Nature* **394,** 299–302
8. Gourley, D. G., Shrive, A. K., Polikarpov, I., Krell, T., Coggins, J. R., Hawkins, A. R., Isaacs, N. W., and Sawyer, L. (1999) *Nat. Struct. Biol.* **6,** 521–525
9. Romanowski, M. J., and Burley, S. K. (2002) *Proteins* **47,** 558–562
10. Krell, T., Coggins, J. R., and Lapthorn, A. J. (1998) *J. Mol. Biol.* **278,** 983–997
11. Schönbrunn, E., Eschenburg, S., Shuttleworth, W. A., Schloss, J. V., Amrhein, N., Evans, J. N., and Kabsch, W. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 1376–1380
12. Chaudhuri, S., and Coggins, J. R. (1985) *Biochem. J.* **226,** 217–223
13. Anton, I. A., and Coggins, J. R. (1988) *Biochem. J.* **249,** 319–326
14. Deka, R. K., Anton, I. A., Dunbar, B., and Coggins, J. R. (1994) *FEBS Lett.* **349,** 397–402
15. Lambert, J. M., Boocock, M. R., and Coggins, J. R. (1985) *Biochem. J.* **226,** 817–829
16. Hawkins, A. R., Lamb, H. K., Moore, J. D., Charles, I. G., and Roberts, C. F. (1993) *J. Gen. Microbiol.* **139,** 2891–2899
17. Onrston, L. N., and Neidle, E. L. (2003) in *The Biology of Acinetobacter* (Tower, K., Bergogne-Berezin, E., and Fewson, C. A., eds) pp. 201–237, Plenum Publishing Corp., New York
18. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
19. Maclean, J., Campbell, S. A., Pollock, K., Chackrewarthy, S., Coggins, J. R., and Lapthorn, A. J. (2000) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56,** 512–515
20. Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) *EMBO J.* **9,** 1665–1672
21. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
22. Collaborative Computational Project, No. 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–763
23. Sheldrick, G. M. (1991) in *Isomorphous Replacement and Anomalous Scattering* (Wolf, W., Evans, P. R., and Leslie, A. G., eds) pp. 80–86, SERC Daresbury Laboratory, Warrington, UK
24. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53,** 240–255
25. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463
26. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
27. 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) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
28. Jones, T. A., Zhou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
29. Geever, R. F., Huiet, L., Baum, J. A., Tyler, B. M., Patel, V. B., Rutledge, B. J., Case, M. E., and Giles, N. H. (1989) *J. Mol. Biol.* **207,** 15–34
30. Hawkins, A. R., Lamb, H. K., Smith, M., Keyte, J. W., and Roberts, C. F. (1988) *Mol. Gen. Genet.* **214,** 224–231
31. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233,** 123–138
32. Conte, L. L., Chothia, C., and Janin, J. (1999) *J. Mol. Biol.* **285,** 2177–2198
33. Walkinshaw, M. D., Taylor, P., Sturrock, S. S., Atanasiu, C., Berge, T., Henderson, R. M., Edwardson, J. M., and Dryden, D. T. (2002) *Mol. Cell* **9,** 187–194
34. Carugo, O., and Argos, P. (1997) *Proteins* **28,** 10–28
35. Chakrabarti, P., and Samanta, U. (1995) *J. Mol. Biol.* **251,** 9–14
36. Dansette, P., and Azerad, R. (1974) *Biochimie* (*Paris*) **56,** 751–755
37. Roszak, A. W., Robinson, D. A., Krell, T., Hunter, I. S., Fredrickson, M., Abell, C., Coggins, J. R., and Lapthorn, A. J. (2002) *Structure* **10,** 493–503
38. Balinsky, D., and Davies, D. D. (1961) *Biochem. J.* **80,** 296–300
39. Bugg, T. D., Abell, C., and Coggins, J. R. (1988) *Tetrahedron Lett.* **29,** 6779–6782
40. Fersht, A. R. (1988) *Biochemistry* **27,** 1577–1580
41. Dennis, A. W., and Balinsky, D. (1972) *Int. J. Biochem.* **9,** 93–102
42. Adams, M. J., Buehner, M., Chandrasekhar, K., Ford, G. C., Hackert, M. L., Liljas, A., Rossmann, M. G., Smiley, I. E., Allison, W. S., Everse, J., Kaplan, N. O., and Taylor, S. S. (1973) *Proc. Natl. Acad. Sci. U. S. A.* **70,** 1968–1972
43. Andersson, P., Kvassman, J., Lindstrom, A., Olden, B., and Pettersson, G. (1981) *Eur. J. Biochem.* **113,** 425–433
44. Krell, T., Chackrewarthy, S., Pitt, A. R., Elwell, A., and Coggins, J. R. (1998) *J. Pept. Res.* **51,** 201–209
45. Chackrewarthy, S. (1996) *Studies on the Active-site of Escherichia coli Shikimate Dehydrogenase.* Ph.D. thesis, University of Glasgow
46. Salgado, H., Santos-Zavaleta, A., Gama-Castro, S., Millan-Zarate, D., Diaz-Peredo, E., Sanchez-Solano, F., Perez-Rueda, E., Bonavides-Martinez, C., and Collado-Vides, J. (2001) *Nucleic Acids Res.* **29,** 72–74
47. Carson, M. (1997) *Methods Enzymol.* **277,** 493–505
48. Gouet, P., Courcelle, E., Stewart, D. I., and Metoz, F. (1999) *Bioinformatics* **15,** 305–308
49. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
50. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins Struct. Funct. Genet.* **11,** 281–296