Crystal Structure of 3-Amino-5-hydroxybenzoic Acid (AHBA) Synthase[†,‡]

Janina C. Eads,*[,§] Morgan Beeby,[§] Giovanna Scapin,[||,⊥] Tin-Wein Yu,[#] and Heinz G. Floss*[,#]

School of Biochemistry, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K., Department of Biochemistry, Albert Einstein College of Medicine, Yeshiva University, Bronx, New York 10461, and Department of Chemistry, Box 35 17 00, University of Washington, Seattle, Washington 98195-1700

Received January 5, 1999; Revised Manuscript Received April 2, 1999

ABSTRACT: The biosynthesis of ansamycin antibiotics, including rifamycin B, involves the synthesis of an aromatic precursor, 3-amino-5-hydroxybenzoic acid (AHBA), which serves as starter for the assembly of the antibiotics' polyketide backbone. The terminal enzyme of AHBA formation, AHBA synthase, is a dimeric, pyridoxal 5′-phosphate (PLP) dependent enzyme with pronounced sequence homology to a number of PLP enzymes involved in the biosynthesis of antibiotic sugar moieties. The structure of AHBA synthase from *Amycolatopsis mediterranei* has been determined to 2.0 Å resolution, with bound cofactor, PLP, and in a complex with PLP and an inhibitor (gabaculine). The overall fold of AHBA synthase is similar to that of the aspartate aminotransferase family of PLP-dependent enzymes, with a large domain containing a seven-stranded β-sheet surrounded by α-helices and a smaller domain consisting of a four-stranded antiparallel β-sheet and four α-helices. The uninhibited form of the enzyme shows the cofactor covalently linked to Lys188 in an internal aldimine linkage. On binding the inhibitor, gabaculine, the internal aldimine linkage is broken, and a covalent bond is observed between the cofactor and inhibitor. The active site is composed of residues from two subunits of AHBA synthase, indicating that AHBA synthase is active as a dimer.

In recent years, genetic approaches to understanding antibiotic biosynthesis have yielded a large amount of information on a number of key biosynthetic pathways (*1*). This has led to the real possibility of engineering new biologically active compounds by genetic manipulation of antibiotic biosynthesis (*2*). The ansamycins are a prominent class of antibiotics produced by various actinomycetes. Because of its pronounced anti-mycobacterial activity rifamycin B (**1**) is one of the more notable members of this class of antibiotics. Synthetically modified versions, e.g., rifampicin, are commonly used in the treatment of tuberculosis and other mycobacterial infections (*3*). However, as with many other antibiotics, the clinical utility of the rifamycins is jeopardized by the widespread appearance of resistant pathogens, calling for the development of new rifamycins not subject to these resistance mechanisms. This has prompted work on the genetics of rifamycin biosynthesis, which recently resulted in the sequencing of the entire rifamycin biosynthetic gene cluster (*4*, *5*). This work laid the ground for a deeper understanding of rifamycin biosynthesis and for the genetic engineering of unnatural analogues of the natural rifamycins.

Like all the ansamycins, rifamycin B is assembled on a type I polyketide synthase by chain extension of an unusual starter unit, the aromatic amino acid 3-amino-5-hydroxybenzoic acid (AHBA)[1] (*6*, *7*), using two acetate and eight propionate units. The carboxy terminus of the resulting linear polyketide eventually forms an amide linkage to the AHBA amino group, thereby closing the macrolactam ring (*8*) (Figure 1). The starter unit AHBA is formed by a new parallel branch *(9)* of the shikimate pathway (*10*) which produces 5-deoxy-5-amino-3-dehydroshikimic acid (aminoDHS), from which AHBA is then formed by the action of the enzyme AHBA synthase (*11*, *12*). Inactivation of the AHBA synthase gene results in the loss of rifamycin B formation; supplementing the mutant with AHBA restores **1** production.

AHBA synthase is a pyridoxal 5′-phosphate (PLP) dependent enzyme which catalyzes the α,β-dehydration and stereospecific 1,4-enolization of aminoDHS to AHBA by the mechanism shown in Figure 2 (*12*). The native enzyme exists in solution as a dimer; its action is irreversibly inhibited by gabaculine (3-amino-2,3-dihydrobenzoic acid), a naturally occurring substrate analogue (*12*). AHBA synthase shows pronounced homology to the deduced amino acid sequences of a group of PLP-dependent enzymes with putative functions as dehydratases or transaminases in sugar nucleotide transformations (*13*, *14*) encoded by various genes cloned from antibiotic biosynthesis gene clusters (cf. ref *12*). Only one

[†] This work was supported by grants from the Royal Society (to J.C.E.) and from the U.S. Public Health Service (NIH Grant AI 20264 to H.G.F.).

[‡] The atomic coordinates of AHBA synthase with PLP and PLP−gabaculine are available from the Brookhaven Protein Data Bank under entry codes 1B9H and 1B9I, respectively.

* To whom correspondence should be addressed.

[§] University of Birmingham.

[||] Yeshiva University.

[⊥] Present address: Merck Research Laboratory, P.O. Box 2000, RY50-105, Rahway, NJ 07065.

[#] University of Washington.

[1] Abbreviations: AHBA, 3-amino-5-hydroxybenzoic acid; aminoDHS, 5-deoxy-5-amino-3-dehydroshikimic acid; AAT, aspartate aminotransferase; AroAT, aromatic amino acid aminotransferase; CBL, cystathionine β-lyase; FOM, figure of merit; Hepes, *N*-(2-hydroxyethyl)-piperazine-*N*′-2-ethanesulfonic acid; MAD, multiple wavelength anomalous diffraction; MIR, multiple isomorphous replacement; PEG, poly(ethylene glycol); PLP, pyridoxal 5′-phosphate; TPL, tyrosine phenol-lyase; Tnase, tryptophanase. In the text an asterisk indicates a residue from the neighboring subunit of the catalytically active dimer.

10.1021/bi990018q CCC: $18.00
Published on Web 07/15/1999

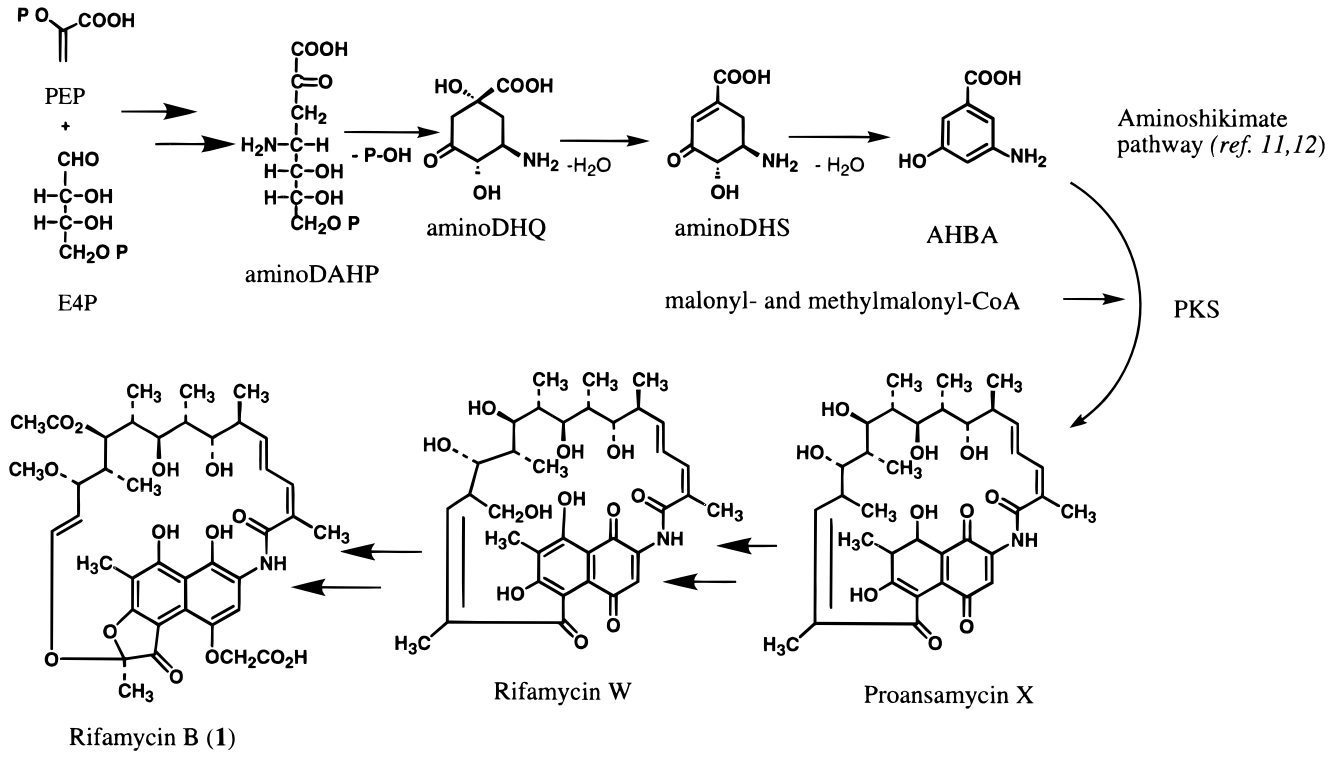

FIGURE 1: Structure of rifamycin B (**1**) and its biosynthetic assembly. P represents the phosphate group in the structures of E4P, PEP, and aminoDAHP.

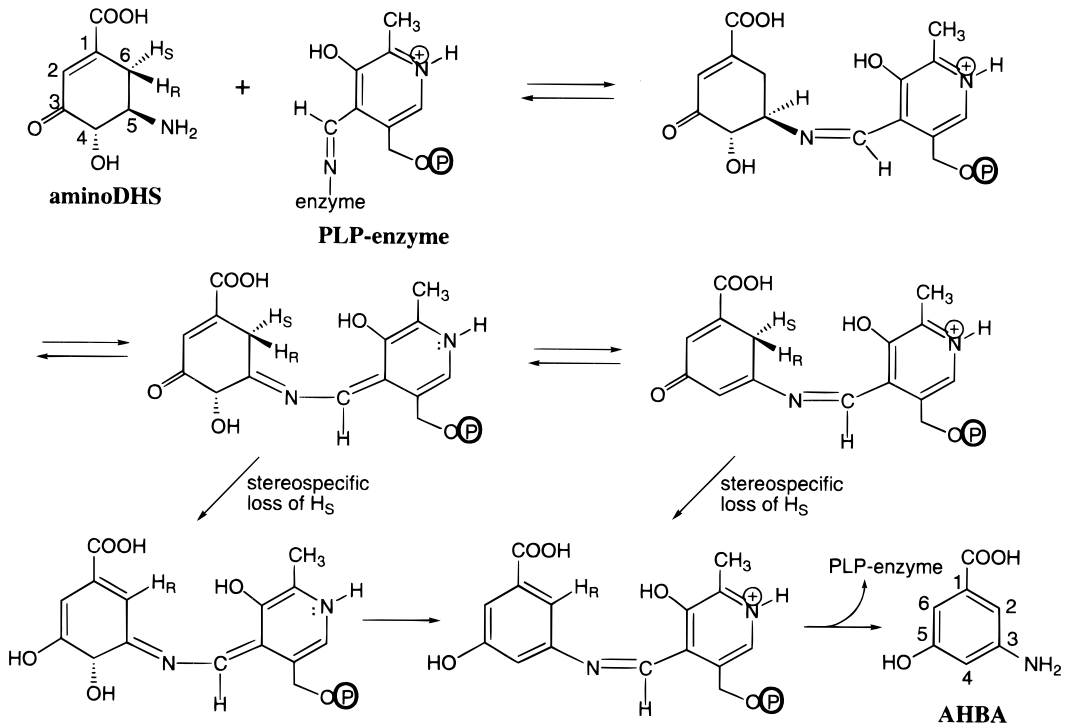

FIGURE 2: Representation of the overall reaction catalyzed by AHBA synthase and its mechanism, with substrate and product numbering as used in the text. (From ref *12*.)

other gene from this family, *stsC* from the streptomycin biosynthetic gene cluster, has so far been expressed, and its product has been characterized as a L-glutamine:*scyllo*-inosose aminotransferase (*15*).

Pyridoxal 5′-phosphate-dependent enzymes catalyze a diverse range of reactions, including transamination, decarboxylation, β-replacement, α,β-elimination, and β,γ-eliminations (*16*, *17*). Recently, the number and variety of PLP-dependent enzyme structures have expanded significantly, allowing careful comparisons of structure−function relationships in this important class of enzymes. Structural studies of enzymes that catalyze different types of reactions will deepen our understanding of how the protein can modulate the chemistry facilitated by the PLP cofactor.

In this paper we present the crystal structures of AHBA synthase with bound PLP and with PLP and the substrate

analogue inhibitor, gabaculine. These structures show that, in terms of three-dimensional structure, AHBA synthase belongs to the aspartate aminotransferase family of PLP-dependent enzymes. The structures also confirm several aspects of the proposed mechanism, including the function of conserved active site residues. They lay the foundation for future studies on the detailed catalytic mechanism of AHBA synthase and for comparative studies with other members of the subfamily of PLP enzymes involved in biosynthesis of antibiotic sugar moieties, to address questions such as what structural features determine whether a given enzyme acts as a dehydratase or as a transaminase.

EXPERIMENTAL PROCEDURES

Overexpression and Purification of Recombinant AHBA Synthase. To express the intact AHBA synthase gene from *Amycolatopsis mediterranei* S699, the 129 bp *Nde*I−*Eco*RI DNA fragment from pRSETB/AHBA *(12)* was replaced with a 24 bp *Nde*I−*Eco*RI digested DNA fragment, which was amplified by PCR using pfu-polymerase (Stratagene), pSK-/AHBA1 *(12)* as template, and the following primers: (i) primer RM11 which carried an *Nde*I site (shown in bold letters) at the 5′-terminus of the AHBA synthase gene (underlined) [5′d(GATGGATCCA**CATATG**AACGCGC-GAAGGCACCGG)3′], containing the ATG start codon of AHBA synthase, and (ii) primer RM16 [5′d(CACCGGCAT-GATCACCTTGGTGCGC)3′], containing the antisense-strand nucleotide sequences located 374 bp downstream of the N-terminus of the AHBA synthase gene. The resulting recombinant clone, pHGF7406, was confirmed through DNA sequencing and then transformed into *Escherichia coli* BL21-(DE3)/pLysS for protein expression (*12*).

A 200-mL overnight culture of *E. coli* BL21(DE3)/pLysS/pHGF7406 in LBBS medium (10 g of tryptone; 5 g of yeast extract; 10 g of NaCl; 182.2 g of D-sorbitol; 0.293 g of betaine per liter) containing 50 μg/mL carbenicillin and 34 μg/mL chloramphenicol was transferred to 10 L of the same fresh medium in a fermentor. Fermentation, harvesting, and washing of the cells were carried out as described previously (*12*). The washed cells were suspended in 100 mL of buffer A (50 mM Tris-HCl, pH 7.5; 10% glycerol; 50 mM KCl; 1 mM EDTA; 1 mM phenylmethanesulfonyl fluoride; 1 mM dithiothreitol). The cell suspension was passed through a French pressure cell twice and centrifuged to separate the cell debris and supernatant. $(NH_4)_2SO_4$ was added to the supernatant, and the precipitate obtained between 58% and 84% saturation was collected, dissolved in buffer B (50 mM Tris-HCl, pH 7.5; 20% glycerol; 50 mM KCl; 1 mM EDTA; 1 mM phenylmethanesulfonyl fluoride), and dialyzed against the same buffer. The dialyzed protein solution was loaded onto a column of DEAE resin (bed volume 185 mL), preequilibrated with buffer B, at a flow rate of 1 mL/min. After being washed with 2 volumes of buffer B, the column was eluted with a gradient of 50−350 mM KCl in buffer B. The fractions containing AHBA synthase were collected (at 200−240 mM KCl) and concentrated using Centriprep-10 (Amicon). $(NH_4)_2SO_4$ was added to the concentrated AHBA synthase solution to give 50% saturation. The solution was applied to a Superose 12 column (bed volume 100 mL), preequilibrated with 50% saturated $(NH_4)_2SO_4$ in buffer B. The column was flushed with buffer B containing 50% saturated $(NH_4)_2SO_4$ for 30 min, followed by elution with a linear gradient of 50−0% saturated $(NH_4)_2SO_4$ in buffer B. Fractions were collected every 10 min, and the fractions containing high AHBA synthase activity, eluting at 24−28% $(NH_4)_2SO_4$ saturation, were monitored for purity by SDS−PAGE and showed the protein to be nearly homogeneous. The AHBA synthase solution from the Superose 12 column, concentrated to a volume of 3 mL in buffer B, was applied to a Sephadex G-100 column (2.5 × 50 cm), which was then eluted with the same buffer at a flow rate of 10 mL/h. Fractions with a high specific activity were collected and were shown to be highly homogeneous, as judged by SDS−PAGE. The overall purification was 74-fold; 435 mg of homogeneous protein was obtained from 10 L of *E. coli* BL21(DE3)/pLysS/pHGF7406 culture.

To obtain partial amino acid sequence data for the expressed recombinant AHBA synthase, the single protein band from the Sephadex G-100 purification step was cut out from the SDS−PAGE gel and electrophoretically transferred to PVDF membrane for gas-phase microsequencing. The intact protein revealed a 14 amino acid N-terminal sequence H_2N-N-A-R-K-A-P-E-F-P-A-W-P-Q-Y, which was consistent with the AHBA synthase sequence obtained directly from *A. mediterranei* S699 (*12*). ESI-MS analysis was carried out on the $NaBH_4$-reduced and the unreduced protein (*12*) and gave a molecular mass of 42 077.0 Da for the denatured apoenzyme (the calculated average molecular mass is 42 039.0 Da) and 42 308.5 Da for the reduced, denatured holoenzyme. The difference between holo and apo forms, 231.5 Da, is consistent with the molecular mass of the phosphopyridoxyl moiety (231 Da).

Crystallization and Data Collection. All chemicals used for crystallization were obtained from Fluka, with the exception of PLP and gabaculine, which were obtained from Sigma. Cryoloops and crystallization supplies were obtained from Hampton Research. All pH measurements were performed at room temperature. Crystals of AHBA synthase were produced by the hanging-drop vapor diffusion method. AHBA synthase (11 mg/mL) in buffer (50 mM HEPES, pH 7.5; 0.05% NaN_3) was incubated at 4 °C with an excess of PLP (0.5 mM) and then mixed 1:1 with a precipitant solution containing 100 mM HEPES, pH 7.5, 1.7 M $(NH_4)SO_4$, and 3−8% PEG 400. The mixture was suspended over 1 mL of the same precipitant solution and left at 18 °C. For crystals containing the inhibitor gabaculine, gabaculine was added (to 0.5 mM) after addition of PLP. The gabaculine−AHBA crystals were almost colorless, whereas the uninhibited form of the enzyme with PLP gave yellow crystals, in accordance with the solution absorption maximum at 416 nm. Crystals of a mercuric acetate derivative of AHBA synthase were produced by cocrystallization of native protein with a 3-fold molar excess of mercuric acetate.

For all data collections, crystals were transferred to cryoprotectant mother liquor (20% v/v glycerol) and then were frozen at 130−140 K in a liquid N_2 cryostream. Diffraction data for crystals of native and gabaculine-inhibited AHBA synthase were collected using a rotating anode generator (Rigaku RU-H2R) and Raxis-IIC image plate detector. A higher resolution dataset for the gabaculine complex was obtained using synchrotron radiation (Daresbury beamline 7.2) with a MAR image plate detector. The diffraction pattern was consistent with the trigonal space

Table 1: Data Collection and Phasing Statistics

	native[a]	Hg λ1	Hg λ2	Hg λ3	gabaculine[b]
λ (Å)	1.54	1.54	1.008	0.832	1.54 and 1.488
resolution range	39−2.35	49−2.0	99−2.4599−1.7	42.3−2.0	
no. of unique reflections	21993	43340	22161	65659	38950
$\langle I/\sigma I\rangle$	7.7 (3.1)[e]	8.5 (2.9)	24.6 (12.0)	18.6 (2.6)	5.1 (4.0)
R_{sym} (%)[c]	7.2 (23.2)	5.7 (24)	3.1 (5.0)	3.7 (20.3)	8.9 (17.7)
completeness	86.7 (58.5)	98 (92.6)	98.9 (99.4)	99.6 (100)	96.3 (89.0)
phasing power[d] (iso)		3.3	3.0	2.5	
phasing power[d] (ano)		2.5	2.5	3.7	
R_{Cullis}[d] (iso)		53	60	66	
R_{Cullis}[d] (ano)		75	76	59	

[a] Native data used for phase calculations, but F_P values calculated by SHARP were used for refinement and map calculation. [b] Data from two crystals were merged for final dataset. [c] $R_{sym} = \sum|(I - \langle I\rangle)|/\sum I$, where I is the observed intensity of a reflection and $\langle I\rangle$ is the mean intensity of the reflection. For the gabaculine dataset, the statistics reflect the agreement between scaled intensities from two datasets. [d] Phasing power and R_{Cullis} given for acentric reflections only. [e] Values in parentheses indicate the values obtained for the highest resolution shell of data, which is the last 0.11 Å of data except for dataset Hg λ3, where it corresponds to the last 0.06 Å of data.

group $P3_121$ or $P3_221$, with unit cell dimensions $a = b =$ 89.7 Å and $c = 127.7$ Å (native data). The asymmetric unit contains one monomer of AHBA synthase (i.e., 387 amino acids, residues 2−388) and 60% solvent. Data were processed with MOSFLM (*18*) and SCALA (*19*), and the statistics for all the merged data sets are given in Table 1. One dataset of the mercuric acetate derivative was collected using a rotating anode generator (Rigaku RU-2HB) equipped with Yale focusing mirrors (Molecular Structures Corp.) and an Raxis-IV image plate detector. Two subsequent datasets of the mercuric acetate derivative were obtained using synchrotron radiation, at 0.832 and 1.008 Å, to coincide with the maximum of f'' above the Hg L(I) edge and the maximum of $|f'|$ at the L(III) edge, respectively. These data were collected with the BM14 beamline at ESRF, Grenoble, using a MarResearch image plate detector and were processed with DENZO and SCALEPACK *(20)*. The data collection statistics are listed in Table 1.

Structure Determination and Refinement. The two mercury sites were located using anomalous difference Patterson maps and confirmed by direct methods (*21*). The site information was then input to SHARP (*22*) for phase calculation and refinement, using data from all three wavelengths for the Hg derivative and the isomorphous native data in a "MAD-MIR" approach to phasing (*23*, *24*). The calculated phases gave excellent FOM values to 1.9 Å (for centric reflections FOM = 0.54; for acentric reflections FOM = 0.58). The resulting phases were further improved by density modification (Solomon solvent flattening), as implemented in SHARP. The final choice of spacegroup was made by inspection of the electron density map: in the correct spacegroup ($P3_221$), the α-helices could be seen to have the correct handedness. The experimental electron density map produced was easily interpreted, showing clear continuous electron density for the entire polypeptide chain from residue 5 and clear electron density for the bound PLP cofactor (see Figure 3). A complete Cα trace (residues 5−388) was built using the automated fitting routine (xfit) in XtalView (*25*). The AHBA synthase sequence could easily be aligned to the Cα trace by inspection of the electron density maps. The backbone atoms and most of the side chains were added to the initial model, using XtalView/Xfit to refine the side chains into the experimental electron density. The initial model produced was subjected to simulated annealing from 3000 K and energy minimization refinement, followed by overall *B*-factor

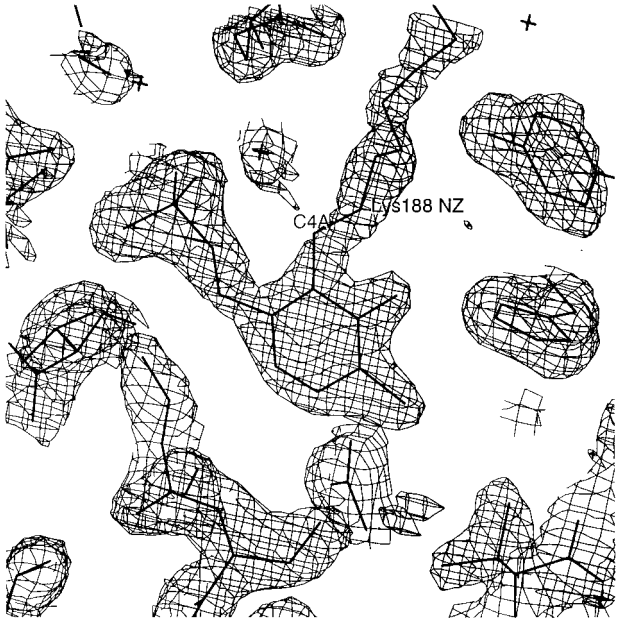

FIGURE 3: Initial electron density map ($|F_P|,\alpha_{MAD-MIR}$) of the AHBA synthase complex with PLP showing the active site area. F_P values and phases were determined from SHARP calculations (see Experimental Procedures). The map is contoured at the 1.5σ level, and the final model is superimposed. C4A is C4′ in the text. The figure was prepared using XtalView (*25*).

refinement as implemented in X-PLOR (*26*), resulting in a model with an *R*-factor of 32.2% and $R_{free} = 35.1\%$ (using data between 10 and 2 Å, $F \geq 1\sigma$). The *F* values used were those calculated by SHARP, as these extended to higher resolution than the native data. Phase restraints were found to improve the refinement at this stage. The PLP cofactor was positioned into the experimental electron density and was added to the model after the first round of refinement. Further model building, followed by energy minimization refinement (X-PLOR, without phase restraints), reduced the *R*-factor to 30.0% ($R_{free} = 32.3\%$). At this point, individual temperature factor refinement was employed, which reduced the *R*-factor to 27.7% ($R_{free} = 30.5\%$). SigmaA-weighted *(27)* electron density maps (calculated in XtalView) were used to aid model building. Further cycles of manual model building, including the addition of water molecules, with positional and temperature factor refinement lowered the *R*-factor to 23.2% ($R_{free} = 26.2\%$). For the last cycles of model building and solvent addition, the structure was refined

Table 2: Statistics for the Refinement of AHBA Synthase Structures

	PLP complex	PLP−gabaculine
resolution range (Å)	10−2.0	20−2.0
R-factor (%)[a]	21.5	22.3
R_{free} (%)	25.2	27.2
no. of reflections in refinement	37744	36758
no. of reflections used for R_{free}	1994	1936
no. of protein atoms	2940	2935
no. of coenzyme and inhibitor atoms	15	24
no. of water molecules	216	190
av *B*-factor for protein atoms (Å^2)	33.7	36.5
av *B*-factor for ligand atoms (Å^2)	25.0	26.4
av *B*-factor for water atoms (Å^2)	39.5	38.8
rms deviations from ideality		
bond lengths (Å)	0.015	0.016
bond angles (deg)	2.3	2.2

[a] The *R*-factor is calculated for all reflections (no cutoffs were applied to the data).

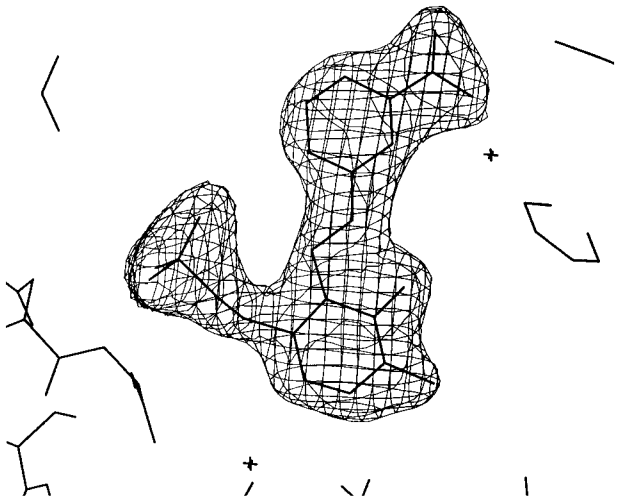

FIGURE 4: Electron density omit map (($|F_o| - |F_c|$),$\alpha_{calc,omit}$) of AHBA synthase complexed with PLP and gabaculine showing the active site area. The substrate and cofactor atoms were omitted from the model after refinement, and the calculation of phases and the omit map were performed in XtalView. The map is contoured at the 6σ level, with the final model superimposed. The figure was prepared using XtalView (*25*).

using the maximum likelihood method (*28*), producing a final model with an *R*-factor of 21.5% (R_{free} = 25.2%). The stereochemistry of the model was assessed with PROCHECK (*29*) (see Table 2).

Structure Determination of the AHBA Synthase−Gabaculine Complex. The refined model of AHBA synthase with PLP, but without added water molecules, was used as the initial model for the structure of AHBA synthase with gabaculine. The model was subjected to rigid body refinement (using the same R_{free} flagged reflection set as for the isomorphous native structure), followed by energy minimization and overall B-factor refinement (*26*). This model was used to make sigmaA-weighted electron density maps in XtalView (*25, 27*). The position of the inhibitor could be clearly observed in the electron density maps (Figure 4), and the inhibitor was added to the model, assuming a single bond between C4′ (C4A) of the cofactor and the nitrogen of gabaculine (see also Figure 10 for chemical structure). Cycles of manual model building and addition of ordered water molecules, combined with energy minimization and individual *B*-factor refinement, reduced the *R*-factor to 21.1% (R_{free} = 26.6%), for data with $F > 2\sigma$, 20−2.3 Å. At this point, data to 2.0 Å were merged with the original data set (see Table 1 for statistics on the merged data set), and addition of water molecules followed by maximum likelihood refinement (*28*) produced a final model with an *R*-factor of 22.4% (R_{free} = 27.2%) for all data to 2.0 Å. The statistics for refinement are given in Table 2.

RESULTS AND DISCUSSION

Quality of the Final Model. The refined structure of the AHBA synthase−PLP complex yields an *R*-factor of 21.5% for all data to 2.0 Å (Table 2). The final model contains residues 5−388 of the 388 residues of AHBA synthase, one PLP molecule, and 216 ordered water molecules. The stereochemistry of the final model has been analyzed by PROCHECK (*29*), which gives an overall *G*-factor of −0.26, within the limits expected for a structure refined at this resolution. A Ramachandran plot shows all non-glycine residues fall within allowed regions of conformational energy. There is one *cis*-proline residue (Pro13), and there are no disulfide bridges. The structure of the enzyme with gabaculine resulted in an *R*-factor of 22.4% for all data to 2.0 Å (Table 2), also with good geometry, and no outliers in the Ramachandran plot.

Description of the Overall Fold of AHBA Synthase. The overall fold of AHBA synthase (shown in Figure 5) consists of three distinct regions, an N-terminal region, a "large" cofactor-binding domain, and a "small" (C-terminal) domain. The large and small domains closely resemble the aspartate aminotransferase fold (*30*), which has now been observed in nine other PLP-dependent enzymes (*31−39*) but is not the only fold observed for a PLP-dependent enzyme (*40, 41*). The N-terminus starts with an extended region, which leads into an α-helix (A1). This leads into the large (cofactor-binding) domain, which consists of a central seven-stranded β-sheet in the order agfedbc (*30*), and direction + − + + + + + (i.e., six parallel strands and one antiparallel strand), surrounded by α-helices, with four (A3, A4, A7, and A10) on one face of the sheet and three (A2, A5, and A6) on the other face. The small (C-terminal) domain consists of a four-stranded antiparallel β-sheet (strands ABDC), with α-helices A9, A13, and A14 and the C-terminal end of helix A8 stacked on the side of the sheet that is furthest from the cofactor-binding domain. The two domains are connected by α-helices A8, A11, and A12.

The asymmetric unit of the AHBA synthase crystal contains one subunit of the enzyme. However, a crystallographic 2-fold axis generates a dimer that has many close contacts between subunits, especially at the active site. This is strong evidence that a dimer is essential for the catalytic activity of AHBA synthase. The observation of a functional dimer in the crystal structure is in agreement with the finding that AHBA synthase behaves as a dimer in solution (*12*). The dimer interface has two main features (Figure 6). The first is the long loop (residues 216−238) between helices A7 and A8, which comprises a large part of the dimer interface and also forms part of the active site. This long loop interacts with the following regions of the other subunit: (i) helix A10 and the preceding seven residues, (ii) residues at the N-terminal end of helix A3, and (iii) residue

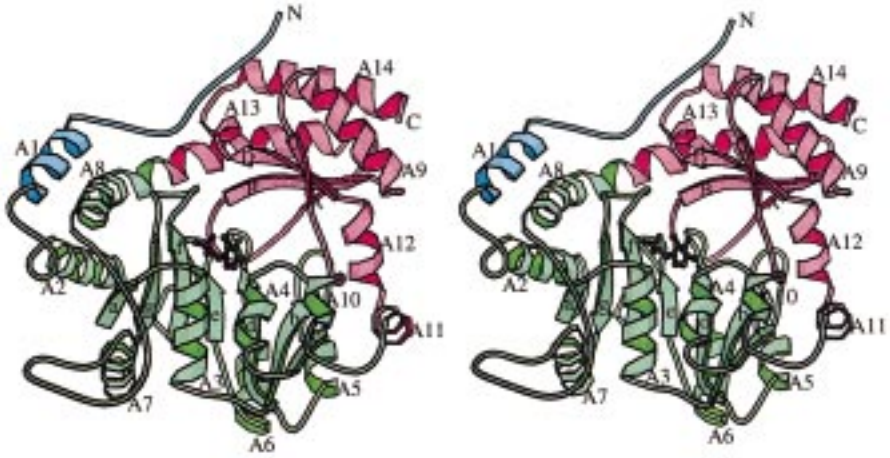

FIGURE 5: Ribbon diagram of the AHBA synthase monomer showing the overall fold. Secondary structure elements are labeled sequentially: A1−A14 for α-helices, a−g for strands in the large domain, and A−D for strands in the small domain. The N-terminal region is colored blue, the large (cofactor-binding) domain is in green, and the small domain is in pink. The cofactor is shown in a ball-and-stick representation. The figure was prepared using Molscript (*49*).

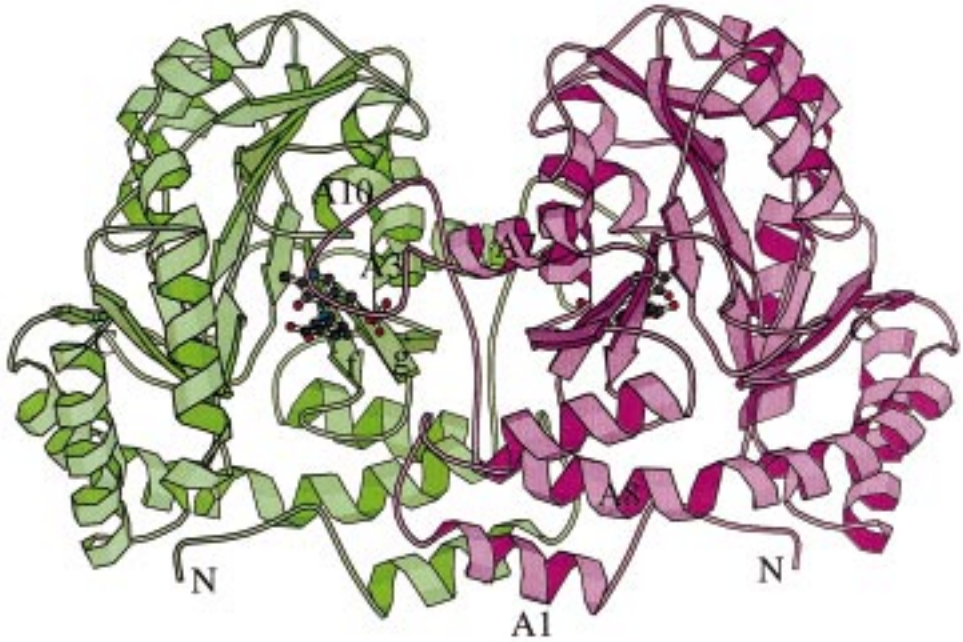

FIGURE 6: Representation of the AHBA synthase dimer. The two monomers are colored green and magenta, with the two molecules of PLP shown in ball-and-stick representation. The figure was prepared using Molscript (*49*).

194 (in the loop between strands f and g). The second major feature of the dimer interface is the N-terminal helix A1 and the following loop up to residue 34. Helix A1 lies antiparallel to its symmetry-related counterpart. The C-terminal residues of helix A1 and the following residues until Arg34 lie near residues of the loop between strands f and g of the second subunit and also near the active site.

Description of the Active Site: PLP Binding. The arrangement of residues at the PLP binding site is shown in Figures 7 and 8. In the PLP form of the enzyme (no inhibitor), the PLP−enzyme forms an internal aldimine species with Lys188. Electron density maps (for example, the initial electron density map shown in Figure 3) show clearly continuous electron density between the side chain amino group (Nζ) of Lys188 and PLP C4′. The refined structure shows that the torsional angle, χ, around the C4−C4′ bond is 33° (χ is the torsional angle defined by the atoms Nζ−C4′−C4−C3). This results in the Lys188 Nζ atom being out of the plane of the pyridoxal ring but still within 2.7 Å of the O3 atom. From the absorption spectrum, which has a maximum at 416 nm, it can be assumed that the internal aldimine is protonated as the ketoenamine tautomer (Figure 7) (*42*), which is consistent with the observed electron density.

The negatively charged O3 atom is also within 2.76 Å of His162. The position of His162 is maintained by a hydrogen bond between the histidine ring nitrogen and amide nitrogen of Ala136. Asp159 forms a salt bridge with the protonated pyridoxal N1. Van der Waals interactions between the cofactor and Ala161, Phe88, and Leu66 also stabilize PLP binding. The torsional angle, ϕ, defined by the atoms OP4−C5′−C5−C4, is 42°, which leaves the phosphate ester oxygen OP4 on the A-face (*si* face) of the cofactor, as observed in CBL (*32*) and Tnase (*33*). This is in contrast to PLP bound to AAT, in which ϕ is −47°, leaving OP4 on the B-face of the cofactor (*30*). The phosphate group of PLP

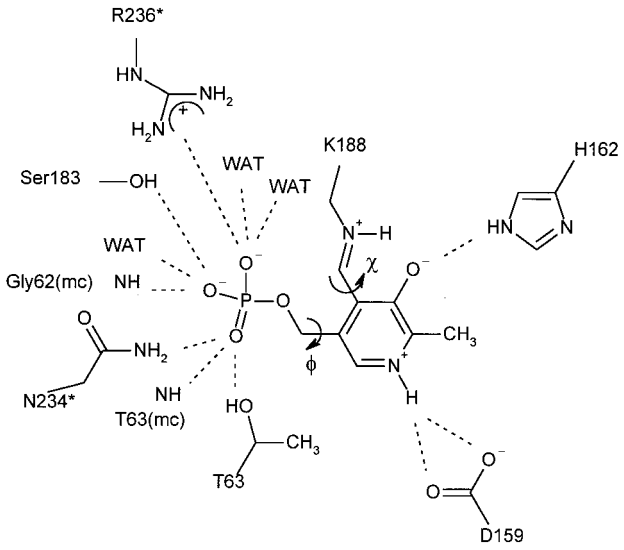

FIGURE 7: Diagram of PLP bound to AHBA synthase showing the residues that form hydrogen bond and charge interactions at the active site. The torsional angles χ and ϕ are marked (see text). WAT denotes a water molecule.

is stabilized by hydrogen bonds and by electrostatic interactions. The phosphate group is positioned at the N-terminal end of α-helix A3 (residues 62−73), providing a positive charge of a helix dipole for stabilization of the phosphate group. In addition, Arg236* is 3.9 Å from the phosphate group and can provide some electrostatic stabilization for the negatively charged phosphate. Hydrogen bonds to the phosphate group are provided by Gly62, Thr63, Ser183, and Asn234* and ordered water molecules. The protonation state of the phosphate group of the cofactor is not known, but it is likely that the interaction with Arg236* and the dipole provided by the N-terminus of helix A3 are sufficient to stabilize the dianionic form at neutral pH.

Description of the Active Site: Inhibitor Form. The binding of gabaculine to AHBA synthase causes small changes at the active site and dimer interface of the enzyme, with the overall tertiary structure remaining essentially unchanged. The rms deviation of Cα positions between the two structures (with and without gabaculine) is 0.2 Å. This is in contrast to the structures of AAT and aromatic amino acid aminotransferase (AroAT), which show significant changes in the small domain on inhibitor binding, reflecting an open and closed form of the enzyme (*31*, *43*). The most significant movement in the AHBA synthase structure on binding gabaculine is a small (ca. 0.6 Å) overall shift in the loop of residues 218*−232*, which is at the dimer interface and also contains residues that form part of the active site. Tyr226* moves to accommodate the gabaculine carboxylate group (the side chain oxygen of Tyr226 is displaced 0.85 Å) and causes the loop to move slightly away from the active site.

When gabaculine binds to AHBA synthase, the characteristic absorption at 416 nm is lost, as the internal aldimine linkage between Lys188 and the cofactor is broken. Consistent with this, there is a clear break in the electron density between Lys188 Nζ and PLP C4′ (not shown). The Lys188 side chain repositions away from the cofactor, leaving the Nζ atom within hydrogen bond distance of the hydroxyl group of Tyr291 and the carbonyl oxygen of Gln185 (see Figure 9). The formation of the covalent bond between the amino nitrogen of gabaculine and C4′ of the cofactor causes only minimal movement of the PLP in its binding pocket—namely, a 10−15° rotation of the plane of the pyridoxal ring, principally around the C5−C5′ bond. The rotation positions C4′ further from Lys188, but leaves N1 and the phosphate group almost unchanged. There are surprisingly few adjustments to the AHBA synthase structure on binding the inhibitor. The carboxylate group of the inhibitor is stabilized by hydrogen bonds to Tyr226* and Arg219* of the second subunit of the catalytic dimer, and as discussed above, Tyr226* is displaced ca. 0.85 Å as gabaculine binds. The aromatic side chain of Phe88 forms an effective stacking interaction with the aromatic ring of the inhibitor (Figure 9). In addition, the Arg236* side chain moves to within 3.4 Å of the PLP phosphate group when gabaculine is bound, allowing a salt bridge interaction (compared to a distance of 3.9 Å in the absence of gabaculine; see Figures 7 and 10). Gabaculine binding displaces several ordered water molecules that are observed at the active site in the structure of the uninhibited form of the enzyme.

Due to the similarity of AHBA and gabaculine, it is reasonable to suggest that the product−enzyme complex would show overall similarity to the gabaculine−enzyme complex, in the orientation of the benzene ring and its amino and carboxylate substituents. In this case, the hydroxyl

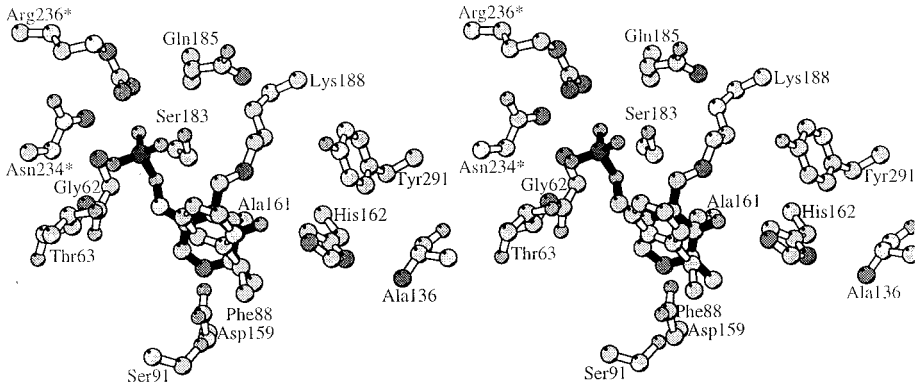

FIGURE 8: Stereo diagram of the active site of AHBA synthase showing the bound cofactor and amino acid side chains at the active site in ball-and-stick representation. Carbon atoms are shown in light gray, oxygen atoms in medium gray, and nitrogen atoms in dark gray, and phosphorus is in black. The bonds of the protein are in white and the bonds of the cofactor are in black. An asterisk indicates a residue from the neighboring subunit of the catalytically active dimer. Prepared using Molscript (*49*).

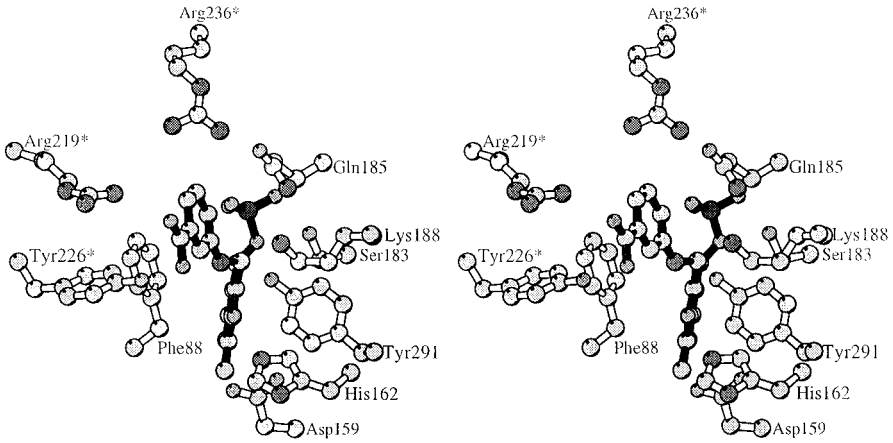

FIGURE 9: Stereo diagram of the active site of AHBA synthase showing the inhibitor gabaculine, PLP, and amino acid side chains at the active site in ball-and-stick representation. The shading of atoms and bonds is as described in Figure 8. For clarity of the gabaculine group, the orientation shown is different from that in Figure 8. An asterisk indicates a residue from the neighboring subunit of the catalytically active dimer.

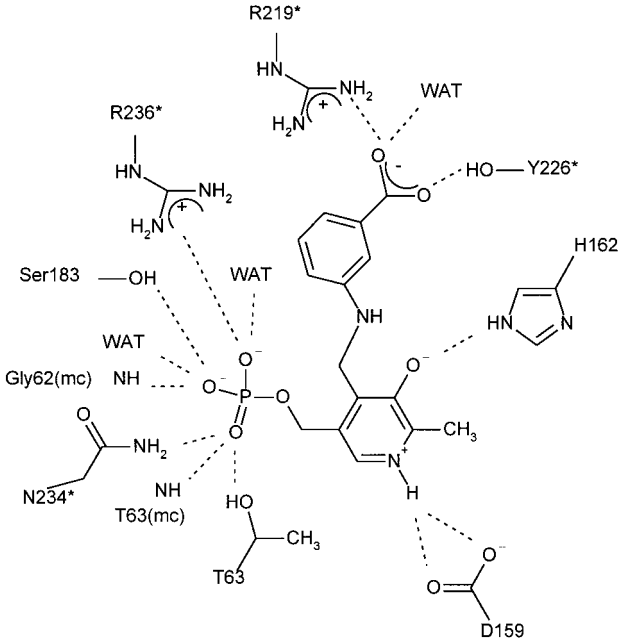

FIGURE 10: Diagram of PLP and gabaculine bound to AHBA synthase showing the residues that provide hydrogen bond interactions at the active site. WAT denotes a water molecule.

substituent of AHBA would be close to the Trp33* ring and the side chain amide group of Gln185.

Comparison of AHBA Synthase with Other PLP-Dependent Enzymes. A search for structurally homologous structures in the Brookhaven database using DALI (*44*) resulted in high similarity scores between AHBA synthase and PLP-dependent enzymes with the aspartate aminotransferase (AAT) fold. There are currently ten enzymes with known 3D structures that show this fold: AAT (*30*), AroAT (*31*), CBL (*32*), Tnase (*33*), tyrosine phenol-lyase (TPL) (*34*), glutamate semialdehyde aminotransferase (*35*), dialkylglycine decarboxylase (*36*), ω-amino acid:pyruvate aminotransferase (*37*), ornithine decarboxylase (*38*), and ornithine aminotransferase (*39*). The cofactor-binding domain is structurally very similar in all these enzymes, with some differences in the length and orientation of the helices and loops connecting the seven strands. When the large domains of the enzymes are overlaid, it is observed that the relative orientations of the small domains differ considerably (*32*). In addition, the small domain varies in the arrangement of β-strands, in the length of loops and helices, and in the involvement of residues from the N-terminus. AHBA synthase has four antiparallel β-strands in the small domain, with no involvement of residues of the N-terminus. AAT has four β-strands in its small domain, but they are arranged somewhat differently, in a 2+2 arrangement, with one antiparallel and one parallel pair of strands and with one strand contributed from the N-terminus. In this aspect, AHBA synthase appears to have the highest homology with CBL.

The enzymes with the AAT fold all exist as dimers, with the dimer interface forming the active site in a generally similar arrangement. However, the contribution of residues to the dimer interface and active site differs considerably between the individual enzymes. This is largely a result of the differences in the topologies of the N-terminal regions and small domains—the overall relationships of the monomers are similar.

There is not a strict conservation of PLP-binding residues in the AAT structural family of PLP-dependent enzymes. The active site lysine (Lys188 in AHBA synthase) is completely conserved, as is the aspartate residue that is close to the PLP N1 (Asp159 in AHBA synthase). The serine located near the PLP phosphate (Ser183 in AHBA synthase) is conserved in nearly all of the AAT subfamily, but interestingly Ser183 is five (rather than three) residues before the active site lysine. This five-residue spacing between Ser and Lys in the PLP-binding site is conserved within the family of putative aminotransferases and dehydratases involved in the biosynthesis of antibiotic sugar moieties, which show strong sequence homology to AHBA synthase.

Many of the remaining active site residues are semiconserved within the AAT subfamily, for example, the stacking of a small side chain (Ala161 in AHBA synthase) and an aromatic side chain (Phe88 in AHBA synthase) on either face of the pyridoxal ring. An interesting lack of conservation is observed near the PLP O3 at the active site, where residues vary from basic (Arg226 in Tnase) to polar (Tyr225 in AAT, His162 in AHBA synthase). This interaction has been proposed to modulate the pK_a of the Schiff base (*45*) and is likely to play a vital role in determining the reaction pathway of AHBA synthase. Among the family of enzymes that show

strong homology to AHBA synthase (see the introduction), the putative transaminases generally have a glutamine or histidine in this position, whereas the putative dehydratases have an acidic residue (*12*−*14*). Thus, as a dehydratase, AHBA synthase appears to be exceptional in having histidine at this position.

Implications for Substrate Binding and Catalysis. The structure of AHBA synthase with the inhibitor gabaculine can be used to predict which groups of the enzyme (and cofactor) are likely to be involved in the binding of substrate and thus may suggest residues that may have roles in the various steps of the proposed catalytic mechanism. Gabaculine (3-amino-2,3-dihydrobenzoic acid) is an analogue of both the substrate and product of AHBA synthase, having a 1,3-arrangement of carboxylate and amino substituents. However, it is possible that the gabaculine complex is not an accurate model of the likely substrate orientation, so too detailed speculation is not warranted. The route from substrate to product involves several steps (Figure 2): (1) formation of the cofactor−substrate external aldimine, (2) abstraction of the α-proton, (3) abstraction of the 4-OH group, (4) abstraction of the *pro*-6*S* proton and protonation of the 3-enolate, and (5) release of the product and formation of the internal aldimine (*12*). The order of steps 3 and 4 has not yet been determined. Stereochemical studies (*12*) together with the orientation of the gabaculine−PLP complex deduced from the crystal structure imply that all enzymatic (or nonenzymatic) steps are carried out on the C-4′ *si* face of the substrate−cofactor complex.

In the AAT and Tnase mechanisms, the "active" form of the internal aldimine is the dipolar ionic form, with a deprotonated imine nitrogen and O3 enolate (*46*, *47*). On formation of the substrate−PLP external aldimine, the active site lysine is in an unprotonated state and is poised to act as the base for α-proton abstraction. In AHBA synthase, the resting state of the enzyme at pH 7.5 appears to be the ketoenamine tautomer of the protonated aldimine, but it is reasonable to suggest that the active form is again the dipolar ionic form, as the His162 residue near O3 would have the effect of lowering the pK_a of the internal aldimine, similar to the effect of Tyr225 and Asn194 in AAT and of Arg226 in Tnase. In CBL, where there is no residue to stabilize the negative charge on O3, the pK_a of the internal aldimine nitrogen is >11, and it has been proposed that the ketoenamine form of the internal aldimine must be the reactive species (*32*). In this case, the active site lysine is released in a fully protonated form and is not able to effect catalysis of α-proton removal. In AHBA synthase, it is proposed that the abstraction of the α-proton is most likely mediated by Lys188, as it is the closest residue and should be released from the internal aldimine in the unprotonated state.

If it is assumed that the carboxylate group of aminoDHS lies in the same position as the carboxylate group of gabaculine and since, as we have proven by the borohydride reduction of the enzyme−substrate complex (*12*), the amino group of aminoDHS forms a external aldimine linkage to C4′ of the cofactor, then the positions of the substrate C4 OH, C6 hydrogen, and C3 carbonyl are fairly restricted. The C4 hydroxyl group, which is eliminated in the reaction, would likely be quite close to the phosphate group of the cofactor and to Lys188. It is possible that Lys188 could donate a proton to the 4-OH group to aid its elimination. However, it is interesting that a phosphate oxygen atom is very close to the likely position of the substrate 4-OH group. The negatively charged phosphate would hinder a negatively charged leaving group, reinforcing the hypothesis that the 4-OH group must be protonated before elimination. An alternative to 4-OH protonation by Lys188 could be envisaged if the phosphate group were protonated (monoanionic form). In this situation, a proton could be donated from the phosphate to 4-OH to effect its elimination. It is likely, however, that the phosphate group is in a dianionic form, as shown for Tnase (*48*), as the presence of the helix dipole and charged group (Asp236) would stabilize the negative charge and lower the pK_a. The 6-H_S abstraction has been shown to be facilitated by the enzyme, as it is completely stereospecific (*12*). Again, Lys188 would be a good candidate base for this step (especially if unprotonated, following protonation of the leaving 4-OH group). Tyr291 is the next closest residue to the 6-H_S position (ca. 4 Å).

In summary, it is probable that the specificity of AHBA synthase for its substrate arises in part from hydrogen bonding between its carboxylate group and Tyr226* and Arg219* and from aromatic stacking between the benzene ring and the Phe88 side chain. The importance of other residues identified at the active site (especially Gln185 and Tyr291) in the reaction mechanism cannot be determined from the structure alone; further experiments will establish their precise roles.

CONCLUSIONS

AHBA synthase shows strong sequence homology to a family of enzymes that have been assigned putative functions as transaminases or dehydratases in sugar nucleotide transformations of antibiotic biosynthesis. The present study predicts that members of this family of enzymes are all likely to contain the familiar L-aspartate aminotransferase fold. The structure of AHBA synthase provides a framework from which to build a detailed understanding of this family of enzymes and to explore what determines whether a particular enzyme acts as a transaminase or dehydratase.

ACKNOWLEDGMENT

The authors thank Dr. Andy Thomson, Dr. Peter Moody, and Mr. Keith Paxton for their help and advice in data collection.

REFERENCES

1. Hopwood, D. A. (1997) *Chem. Rev. 97*, 2465−2497.
2. Khosla, C. (1997) *Chem. Rev. 97*, 2577−2590.
3. Sepkowitz, K. A., Raffali, J., Riley, L., Kiehn, T. E., and Armstrong, D. (1995) *Clin. Microbiol. Rev. 8*, 180−199.
4. August, P. R., Tang, L., Yoon, Y. J., Ning, S., Müller, R., Yu, T.-W., Taylor, M., Hoffmann, D., Kim, C.-G., Zhang, X., Hutchinson, C. R., and Floss, H. G. (1998) *Chem. Biol. 5*, 69−79.
5. Schupp, T., Toupet, C., Engel, N., and Goff, S. (1998) *FEMS Microbiol. Lett. 159*, 201−207.
6. Ghisalba, O., and Nüesch, J. (1981) *J. Antibiot. 34*, 64−71.
7. Anderson, M. G., Monypenny, D., Rickards, R. W., and Rothchild, J. (1989) *J. Chem. Soc., Chem. Commun.*, 311−313.
8. Lancini, G., and Cavalleri, B. (1997) in *Biotechnology of Antibiotics* (Strohl, W. R., Ed.) 2nd ed., pp 521−549, Dekker, New York.
9. Floss, H. G. (1997) *Nat. Prod. Rep. 14*, 433−452.

10. Haslam, E. (1993) *Shikimic Acid: Metabolism and Metabolites*, John Wiley and Sons, Chichester.
11. Kim, C.-G., Kirschning, A., Bergon, P., Zhou, P., Su, E., Sauerbrei, B., Ning, S., Ahn, Y., Breuer, M., Leistner, E., and Floss, H. G. (1996) *J. Am. Chem. Soc. 118*, 7486−7491.
12. Kim, C.-G., Yu, T.-W., Fryhle, C. B., Handa, S., and Floss, H. G. (1998) *J. Biol. Chem. 273*, 6030−6040.
13. Thorson, J. S., Lo, S. F., Liu, H.-w., and Hutchinson, C. R. (1993) *J. Am. Chem. Soc. 115*, 6993−6994.
14. Pascarella, S., and Bossa, F. (1994) *Protein Sci. 3*, 701−705.
15. Ahlert, J., Distler, J., Mansouri, K., and Piepersberg, W. (1997) *Arch. Microbiol. 168*, 102−113.
16. Alexander, F. W., Sandmeier, E., Mehta, K., and Christen, P. (1994) *Eur. J. Biochem. 219*, 953−960.
17. Vederas, J. C., and Floss, H. G. (1980) *Acc. Chem. Res. 13*, 455−463.
18. Leslie, A. G. W. (1992) *CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, Number 26, SERC Daresbury Laboratory, Warrington, U.K.
19. Evans, P. R. (1993) Data reduction, in *Proceedings of CCP4 Study Weekend on Data Collection & Processing*, pp 114−122, SERC Daresbury Laboratory, Warrington, U.K.
20. Otwinowski, Z., and Minor, W. (1993) *DENZO. A film processing program for macromolecular crystallography*, Yale University, New Haven, CT.
21. Sheldrick, G. M. (1993) *Acta Crystallogr. D49*, 18−23.
22. La Fortelle, E. de, Irwin, J. J., and Bricogne, G., (1997) *Crystallographic Computing 7* (Bourne, P., and Watenpaugh, K., Eds.) Oxford Press, Oxford, U.K.
23. Ramakrishnan, V., and Biou, V. (1997) *Methods Enzymol. 276*, 538−557.
24. Terwilliger, T. C. (1997) *Methods Enzymol. 276*, 530−537.
25. McRee, D. E., and Israel, M. (1997) XtalView, in *Crystallographic Computing* (Bourne, E., and Watenpaugh, K., Eds.) Oxford Press, Oxford, U.K.
26. Brunger, A. T. (1992) *X-PLOR, Version 3.1. A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
27. Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
28. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D53*, 240−255.
29. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
30. Ford, G. C., Eichele, G., and Jansonius, J. N. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 2559−2563.
31. Okamoto, A., Nakai, Y., Hayashi, H., Hirotsu, K., and Kagamiyama, H. (1998) *J. Mol. Biol. 280*, 443−461.
32. Clausen, T., Huber, R., Laber, B., Pohlenz, H.-D., and Messerschmidt, A. (1996) *J. Mol. Biol. 262*, 202−224.
33. Isupov, M. N., Antson, A. A., Dodson, E. J., Dodson, G. G., Dementieva, I. S., Zakomirdina, L. N., Wilson, K. S., Dauter, Z., Lebedev, A. A., and Harutyunyan, E. H. (1998) *J. Mol. Biol. 276*, 603−623.
34. Antson, A. A., Demidkina, T. V., Gollnick, P., Dauter, Z., Von Tersch, R. L., Long, J., Berezhnoy, S. N., Phillips, R. S., Harutyunyan, E. H., and Wilson, K. S. (1993) *Biochemistry 32*, 4195−4206.
35. Hennig, M., Grimm, B., Contestabile, R., John, R. A., and Jansonius, J. N. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 4866−4871.
36. Toney, M. D., Hohenester, E. H., Cowan, S. W., and Jansonius, J. N. (1993) *Science 261*, 756−759.
37. Watanabe, N., Sakabe, K., Sakabe, N., Higashi, T., Sasaki, K., Aibara, S., Morita, Y., Yonaha, K., Toyama, S., and Fukutani, H. (1989) *J. Biochem. 105*, 1−3.
38. Momany, C., Ernst, S., Ghosh, R., Chang, N., and Hackert, M. L. (1995) *J. Mol. Biol. 252*, 643−655.
39. Shen, B. W., Hennig, M., Hohenester, E., Jansonius, J. N., and Schirmer, T. (1998) *J. Mol. Biol. 227*, 81−102.
40. Shaw, J. P., Petsko, G. A., and Ringe, D. (1997) *Biochemistry 36*, 1329−1342.
41. Sugio, S., Petsko, G. A., Manning, J. M., Soda, K., and Ringe, D. (1995) *Biochemistry 34*, 9661−9669.
42. Kallen, R. G., Korpela, T., Martell, A. E., Matsushima, Y., Metzler, C. M., Metzler, D. E., Morozov, Y. V., Ralston, I. M., Savin, F. A., Torchinsky, Y. M., and Ueno, H. (1985) in *Transaminases* (Christen, P., and Metzler, D. E., Eds.) Chapter 2, pp 38−108, Wiley, New York.
43. Okamoto, A., Higuchi, T., Hirotsu, K., Kuramitsu, S., and Kagamiyama, H. (1994) *J. Biochem. 116*, 95−107.
44. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−138.
45. Yano, T., Mizuno, T., and Kagamiyama, H. (1993) *Biochemistry 32*, 1810−1815.
46. Arnone, A., Christen, P., Jansonius, J. N., and Metzler, D. E. (1985) in *Transaminases* (Christen, P., and Metzler, D. E., Eds.) Chapter 5, pp 226−362, Wiley, New York.
47. Phillips, R. S. (1991) *Biochemistry 30*, 5927−5934.
48. Schnackerz, K. D., and Snell, E. E. (1983) *J. Biol. Chem. 258*, 4839−4841.
49. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI990018Q