Structural Studies of the l -Threonine -O- 3-phosphate Decarboxylase (CobD) Enzyme from Salmonella enterica : The Apo, Substrate, and Product−Aldimine Complexes † , ‡

ARTICLE *in* BIOCHEMISTRY · AUGUST 2002

Impact Factor: 3.02 · DOI: 10.1021/bi020294w · Source: PubMed

3 AUTHORS, INCLUDING:

Jorge C Escalante-Semerena
University of Georgia
174 PUBLICATIONS **5,060** CITATIONS
SEE PROFILE

Ivan Rayment
University of Wisconsin–Madison
232 PUBLICATIONS **14,628** CITATIONS
SEE PROFILE

Structural Studies of the L-Threonine-*O*-3-phosphate Decarboxylase (CobD) Enzyme from *Salmonella enterica*: The Apo, Substrate, and Product−Aldimine Complexes[†,‡]

Cheom-Gil Cheong,[§] Jorge C. Escalante-Semerena,*[,||] and Ivan Rayment*[,§]

Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, and Department of Bacteriology, University of Wisconsin, Madison, Wisconsin 53706

Received April 19, 2002; Revised Manuscript Received May 21, 2002

ABSTRACT: The evolution of biosynthetic pathways is difficult to reconstruct in hindsight; however, the structures of the enzymes that are involved may provide insight into their development. One enzyme in the cobalamin biosynthetic pathway that appears to have evolved from a protein with different function is L-threonine-*O*-3-phosphate decarboxylase (CobD) from *Salmonella enterica*, which is structurally similar to histidinol phosphate aminotransferase [Cheong, C. G., Bauer, C. B., Brushaber, K. R., Escalante-Semerena, J. C., and Rayment, I. (2002) *Biochemistry 41*, 4798−4808]. This enzyme is responsible for synthesizing (*R*)-1-amino-2-propanol phosphate which is the precursor for the linkage between the nucleotide loop and the corrin ring in cobalamin. To understand the relationship between this decarboxylase and the aspartate aminotransferase family to which it belongs, the structures of CobD in its apo state, the apo state complexed with the substrate, and its product external aldimine complex have been determined at 1.46, 1.8, and 1.8 Å resolution, respectively. These structures show that the enzyme steers the breakdown of the external aldimine toward decarboxylation instead of amino transfer by positioning the carboxylate moiety of the substrate out of the plane of the pyridoxal ring and by placing the α-hydrogen out of reach of the catalytic base provided by the lysine that forms the internal aldimine. It would appear that CobD evolved from a primordial PLP-dependent aminotransferase, where the selection was based on similarities between the stereochemical properties of the substrates rather than preservation of the fate of the external aldimine. These structures provide a sequence signature for distinguishing between L-threonine-*O*-3-phosphate decarboxylase and histidinol phosphate aminotransferases, many of which appear to have been misannotated.

The proteins required for the de novo synthesis of enzymatic cofactors such as hemes or cobalamin represent a substantial component of the genetic makeup of simple prokaryotes because of the chemical complexity of the pathways and the number of enzymes required to achieve total biosynthesis. At this point in time, the benefit that these cofactors provide far outweighs the biosynthetic cost; however, this raises the question of how these pathways evolved, especially since they arose very early in the development of life.

It is widely accepted that enzyme evolution can occur via gene duplication followed by divergence of catalytic function by mutation and rearrangement. If the development of biosynthetic pathways followed such an adventitious route, it should be expected that the three-dimensional structures of the current enzymes would bear the structural signature of the enzyme from which they diverged. Although there are insufficient protein structures available to test this hypothesis, in the biosynthetic pathway for cobalamin, it appears that enzymes that catalyze several unique steps evolved from other pathways.

Cobalamin is a complex cofactor that requires more than 25 committed enzymes for its biosynthesis (Figure 1). The assembly pathway is normally described in three parts, the synthesis of the corrin ring, attachment of the upper 5′-deoxyadenosine ligand to the cobalt ion, and synthesis and assembly of the nucleotide loop (*1*, *2*). Structural studies on the enzymes involved in the last two parts of the pathway have revealed several unanticipated similarities between proteins in the cobalamin biosynthetic pathway and other critically important enzymes. For example, the enzyme responsible for forming the cobalt−carbon bond and the enzyme involved in activating the aminopropanol arm (CobA and CobU in *Salmonella enterica*, respectively) both are structurally similar to the RecA protein which is an essential enzyme in the control of DNA repair, F1ATPase, and the helicase domain of the gene 4 protein from bacteriophage T7 (*3*−*5*).

[†] This research was supported in part by NIH Grants GM58281 to I.R. and GM40313 to J.C.E.-S. Use of Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Basic Energy Research, under Contract W-31-109-ENG-38.

[‡] The X-ray coordinates have been deposited in the Protein Data Bank as entries 1LC5, 1LC7, and 1LC8 for apo-CobD, the apo-CobD·substrate complex, and the cobD·product aldimine complex, respectively.

* To whom correspondence may be addressed. I.R.: Department of Biochemistry, 433 Babcock Dr., Madison, WI 53706; phone, (608) 262-0437; fax, (608) 262-1319; e-mail, Ivan_Rayment@biochem.wisc.edu. J.C.E.-S.: Department of Bacteriology, Fred Hall, Linden Drive, Madison, WI 53706; e-mail, escalante@bact.wisc.edu.

[§] Department of Biochemistry.

[||] Department of Bacteriology.

10.1021/bi020294w CCC: $22.00
Published on Web 06/22/2002

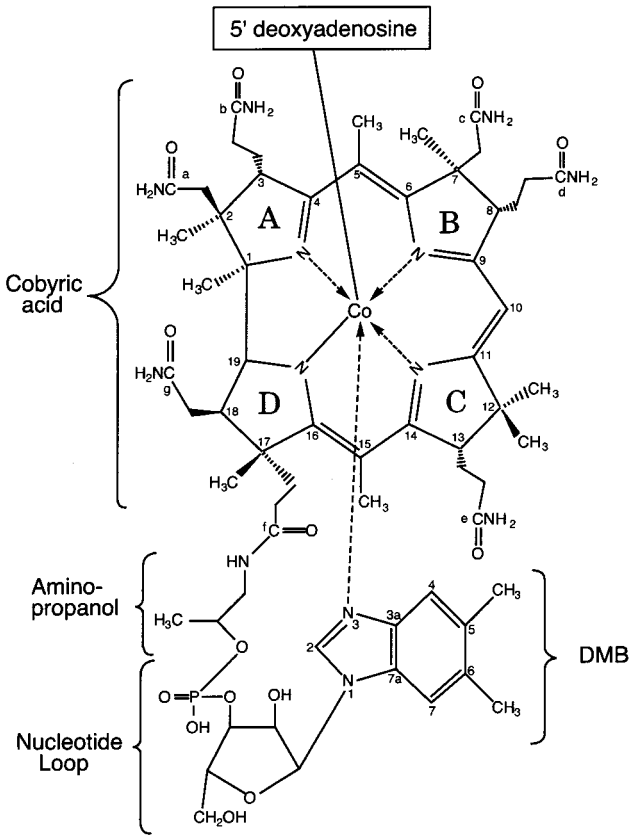
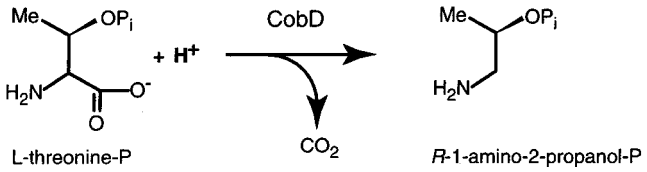


FIGURE 1: Chemical structure of 5′-deoxyadenosylcobalamin. In *S. enterica*, the lower ligand is 5,6-dimethylbenzimidazole.

Scheme 1



Another enzyme that clearly evolved from a different pathway is L-threonine-*O*-3-phosphate decarboxylase (Scheme 1) which is responsible for synthesizing (*R*)-1-amino-2-propanol phosphate, the precursor for the linkage between the nucleotide loop and the corrin ring in cobalamin (*6*). This enzyme was initially believed to be a PLP-dependent aminotransferase based on its sequence being similar to the sequence of this class of enzymes (*7*) but was subsequently shown to be a decarboxylase (*6*). The structure and active site of this enzyme have been shown to be closely related to those of histidinol phosphate aminotransferase (*8*), where the latter is a member of the aspartate aminotransferase subfamily of the α-family of PLP-dependent enzymes or fold type I (*9*, *10*). This decarboxylase activity is unusual since within the aspartate aminotransferase subfamily the aminotransferase activity is highly conserved. The only other member of the aspartate aminotransferase family that does not exhibit aminotransferase activity is 1-aminocyclopropane-1-carboxylate synthase (*11*).

CobD is a comparatively small dimeric protein containing 364 amino acid residues which form two domains. One domain is a typical PLP-binding motif, whereas the second is built from segments at both the N- and C-termini of the protein and contributes to the substrate binding pocket. Unlike most other members of this family, the N-terminal extension is quite short and does not form the same type of interdomain exchange of secondary structural elements common to other aminotransferases. Consequently, it is smaller than many members of the α-family of PLP-dependent enzymes. The interesting feature of CobD is that it appears to have evolved from an aminotransferase rather than an amino acid decarboxylase which is surprising considering that both types of enzyme are found within the extensive α-family of PLP-dependent enzymes (*10*). Although this type of evolutionary development is unusual in the aspartate aminotransferase subgroup, dialkylglycine decarboxylase, which belongs to a different subgroup, can also be viewed as an aminotranferase that has developed a decarboxylase activity, though it has retained its aminotransferase function (*12*).

In both the aminotransferases and decarboxylases, the first chemical step is formation of a Schiff base with PLP via a transimination reaction to form an external aldimine. Thereafter, the fate of this intermediate is stereochemically controlled to yield a wide range of products, including either amino transfer or decarboxylation (*13*, *14*). Thus, the question for CobD is whether it was easier to evolve substrate specificity as would be needed for remodeling of a decarboxylase or reaction specificity required for conversion of an aminotransferase into a decarboxylase. Thus, was it serendipity or design that allowed the evolution of a histidinol aminotransferase into a decarboxylase? The structure of CobD suggested that the chemical fate of the external aldimine could have been redirected most easily by modifications at the N-terminus of an aminotransferase (*8*). The purpose of this paper is to address these questions by determining the structures of CobD in its apo form, complexed with L-threonine phosphate, and the product external aldimine intermediate in an effort to understand those factors that determine specificity and stereochemical preference for a decarboxylation reaction relative to the deamination. These structures are reported at 1.46, 1.8, and 1.8 Å resolution, respectively, and yield insight into nature's choice of retooling an aminotransferase rather than a decarboxylase to accomplish the biosynthesis of 1-amino-2-propanol phosphate.

MATERIALS AND METHODS

Protein Purification. L-Threonine-*O*-3-phosphate decarboxylase was expressed and purified as described previously (*8*). Prior to crystallization, CobD was dialyzed overnight at 4 °C against 50 mM Hepes (pH 7.0), 150 mM NaCl, and 2 mM DTT[1] concentrated to 10 mg/mL. Thereafter, the protein was frozen dropwise in liquid nitrogen and stored at −80 °C for future use.

Preparation of the Apoenzyme. Efforts to inactivate the enzyme by reducing the Schiff base between PLP and Lys^{216} with sodium borohydride inadvertently generated the apoenzyme as follows. Protein at 4.4 mg/mL (∼0.1 mM) was dialyzed overnight against 200 mM sodium borate buffer (pH 8.5) containing 100 mM NaCl prior to reduction. A sodium borohydride solution (100 mM) was prepared at 4

[1] Abbreviations: DTT, dithiothreitol; rms, root-mean-square; MES, 2-(*N*-morpholino)ethanesulfonic acid.

°C in distilled water. An approximately 10-fold excess of a 100 mM stock solution of sodium borohydride dissolved in water was added with stirring to 10 mL of protein solution for 10 min. Immediately on addition of the sodium borohydride, the yellow color of the protein solution disappeared. The progress of the reduction was revealed spectroscopically with a Beckman DU640B spectrophotometer where the UV absorption spectra initially exhibited a peak at 420 nm which shifted to 330 nm within several minutes, indicating reduction of the PLP cofactor. Electrospray mass spectrometry for both the reduced apo-CobD and wild-type CobD revealed that the masses of these two samples were 40 666 and 40 894 Da, respectively. The difference between these samples was consistent with the mass of the PLP moiety of the PLP−lysine adduct (230 Da) less two protons which were present on the lysine side chain after PLP was removed.

Crystallization of Apo-CobD. Reduced apo-CobD was extensively dialyzed at 4 °C against 25 mM HEPES buffer (pH 7.5), 200 mM NaCl, and 5 mM DTT before crystallization. Thereafter, the protein solution was concentrated to 10 mg/mL and flash-frozen as small droplets in liquid nitrogen. Crystals of apo-CobD were grown with the hanging-drop vapor-diffusion technique where equal volumes of protein at 10 mg/mL in its final storage buffer and a precipitant containing 1.1 M sodium potassium phosphate and 50 mM lithium sulfate (pH 6.0) were mixed and suspended over the same precipitant solution at room temperature. Crystals grew spontaneously and achieved dimensions of 0.4 mm × 0.4 mm × 0.4 mm within 1 week. Prior to being frozen, crystals were transferred to a synthetic mother liquor composed of 1.5 M sodium potassium phosphate (pH 6.0). Crystals were cryoprotected by being transferred to the same synthetic mother liquor but also containing 30% glycerol. The crystals belong to space group *I*222 with the following unit cell dimensions: $a = 76.0$ Å, $b = 103.3$ Å, and $c = 109.3$ Å; the crystal lattice contains one monomer per asymmetric unit with a solvent content of 53%.

Crystallization of the Apo-CobD·L-Threonine Phosphate Complex. Crystals of apo-CobD in complex with L-threonine phosphate (substrate complex) were grown with the small-scale batch technique (*15*). L-Threonine phosphate was mixed with a protein solution to a final concentration of 50 mM L-threonine phosphate and 9 mg/mL protein, and the protein solution was incubated at 4 °C for 1 h. Ten microliters of the protein solution and 10 μL of a precipitant solution composed of 10% methyl ether PEG 2000, 100 mM sodium potassium phosphate, and 100 mM MES buffer (pH 6.0) were mixed carefully and sealed in nine-well depression plates at room temperature. Crystals grew without seeding and achieved dimensions of 0.2 mm × 0.2 mm × 0.6 mm within 1 day. Prior to being cryocooled, the crystals were transferred to a synthetic mother liquor consisting of 10% methyl ether PEG 2000, 100 mM NaCl, 50 mM L-threonine phosphate, and 100 mM MES buffer (pH 6.0). The crystals were stable in this solution and soaked for 10 days before being frozen. Crystals were cryoprotected by transfer to a cryoprotectant solution consisting of 25% methyl ether PEG 2000, 1.5 M NaCl, 50 mM L-threonine phosphate, and 100 mM MES buffer (pH 6.0). Crystals belong to space group *I*222 with the following unit cell dimensions: $a = 66.6$ Å, $b = 104.0$ Å, and $c = 117.9$ Å; the crystal lattice contains one monomer per asymmetric unit with a solvent content of 51%.

Table 1: Data Collection Statistics

	apo-CobD	substrate complex	reaction intermediate complex
resolution (Å)	1.46	1.8	1.8
no. of unique reflections	74541	38108	37471
redundancy	8.2	8.8	7.6
completeness (%)[a]	99.4 (95.4)	99.6 (99.1)	99.4 (99.2)
average I/σ	66.7 (13.2)	60.8 (10.2)	31.9 (5.8)
R_{merge}[b]	0.040 (0.129)	0.055 (0.219)	0.078 (0.331)

[a] The numbers in parentheses represent completeness in the highest-resolution shell. [b] $R_{merge} = (\sum|I_{hkl} - I|)/(\sum I_{hkl})$, where the average intensity I is taken over all symmetry equivalent measurements and I_{hkl} is the measured intensity for a given reflection.

Crystallization of the CobD Reaction Intermediate Complex. The crystals employed for the structural investigation of the reaction intermediate complex were grown with the small-scale batch technique (*15*). L-Threonine phosphate was added to wild-type CobD to final concentrations of 25 mM L-threonine phosphate and 8 mg/mL protein, and the protein solution was incubated at 4 °C for 1 h. Ten microliters of the protein solution and 10 μL of a precipitant solution composed of 4% methyl ether PEG 2000, 150 mM NaCl, and 100 mM MES buffer (pH 6.0) were mixed carefully and incubated at room temperature. Crystals grew spontaneously and achieved dimensions of 0.2 mm × 0.2 mm × 0.4 mm within 1 week. Before being frozen, crystals were transferred to a synthetic mother liquor consisting of 4% methyl ether PEG 2000, 100 mM NaCl, and 100 mM MES buffer (pH 6.0). Prior to cryoprotection, the crystals were transferred to a solution consisting of 25% methyl ether PEG 2000, 1.5 M NaCl, and 100 mM MES buffer (pH 6.0). The crystals belonged to space group *I*222 with the following unit cell dimensions: $a = 66.6$ Å, $b = 103.2$ Å, and $c = 117.1$ Å; the crystal lattice contains one monomer per asymmetric unit with a solvent content of 50%.

X-ray Data Collection. X-ray intensities were measured at 100 K on a 3 × 3 charge-coupled device area detector with synchrotron radiation at a wavelength of 0.9763 Å on beam line 19-BM of the Advanced Photon Source (Structural Biology Center-CAT, Argonne National Laboratory, Argonne, IL). The data sets were integrated and scaled with the HKL2000 program suite (*16*). Data collection statistics are summarized in Table 1.

Structural Determination of Apo-CobD. The structure of apo-CobD was determined by molecular replacement with the wild-type CobD structure as a search model (*8*). The program package AMoRe was employed using data between 8.0 and 2.8 Å, which gave a clear peak with a correlation coefficient of 71.3% and an *R*-factor of 38.6%. The structure was initially refined at 2.3 Å resolution with the program CNS (*17*) and manually adjusted with TURBO-FRODO (*18*). Thereafter, the structure was completed with stepwise refinement and manual adjustment at 2.0, 1.7, and 1.46 Å resolution where 5% of reflections were excluded for calculation of the cross-validated *R*-factor (R_{free}). Water molecules were incorporated into the structure if they had peak heights of $>3\sigma$ in the $F_o - F_c$ difference Fourier map and were within hydrogen bonding distance of appropriate atoms. The final model was refined to an *R*-factor of 19.0%

Table 2: Refinement Statistics[a]

	apo-CobD	apo-CobD substrate complex	product aldimine complex
resolution limits (Å)	1.46	1.8	1.8
final *R*-factor[b]	19.0	21.2	19.6
R_{free}	21.4	25.1	22.9
no. of reflections (working set)	70826	35941	35608
no. of reflections (test set)	3715	1908	1862
no. of protein atoms	2775	2733	2735
no. of solvent molecules	271	228	271
other molecules, ions	2 PO_4	L-threonine phosphate, 1 PO_4	product aldimine
average *B* value			
main chain atoms	14.4	25.8	19.7
all protein atoms	16.4	27.4	21.4
solvent atoms	25.4	33.6	28.5
ligand atoms	10.4	30.4	14.6
weighted rms deviations from ideality			
bond lengths (Å)	0.011	0.011	0.013
bond angles (deg)	1.54	1.53	1.57
PDB entry	1LC5	1LC7	1LC8

[a] CNS refinement. [b] $R = \sum||F_o| - k|F_c||/\sum|F_o|$, where 5% of the data were excluded for calculation of R_{free}.

and an R_{free} of 21.4%. Finally, 92.4% of all residues lie in the most favored regions of the Ramachandran plot, and no residues are located in the disallowed region, as determined with the program PROCHECK (*19*).

The structures of the substrate complex (apo-CobD in complex with L-threonine phosphate) and the product aldimine complex were also determined by molecular replacement utilizing refined apo-CobD and the wild-type CobD structure as search models, respectively. The program package AMoRe was employed using data between 8.0 and 2.8 Å resolution, which gave clear peaks for the substrate complex and the reaction intermediate complex with correlation coefficients of 71.1 and 80.2% and *R*-factors of 37.8 and 32.5%, respectively. After molecular refinement, the structures were refined to 1.8 Å resolution with the program CNS and then manually rebuilt with TURBO-FRODO. Five percent of the reflections were excluded from calculation of the cross-validated *R*-factor (R_{free}). Water molecules were incorporated into the structure if they had peak heights of $>3\sigma$ in the $F_o - F_c$ difference Fourier map and were within hydrogen bonding distance of appropriate atoms. The final models were refined to *R*-factors of 21.2 and 19.6% and R_{free} values of 25.1 and 22.9% for substrate and product aldimine complexes, respectively. Finally, 92.4 and 90.8% of all residues lie in the most favored regions of the Ramachandran plot for substrate and product aldimine complexes, respectively, and no residues are located in the disallowed region, as determined with the program PROCHECK (*19*).

RESULTS AND DISCUSSION

Structure of Apo-CobD. The final model for apo-CobD contains 355 amino acid residues out of a total of 364 and includes 271 water molecules and two phosphate ions. The electron density extends continuously from Leu^3 to Leu^{357}. A section of the electron density is shown in Figure 2A. A ribbon representation of the crystallographic subunit is shown in Figure 3 which reveals, as seen previously in the structure of the PLP·CobD complex (*8*), that the molecule consists of a large domain and a small domain. The large domain is a typical PLP-binding domain found in the aspartate aminotransferase family of PLP-dependent enzymes (*9*, *20*). The overall structure is very similar to that of the PLP·CobD complex where the overall rms difference between 320 α-carbon atoms is 0.42 Å. Thus, removal of PLP does not induce any overall conformational changes such as a change in the relative orientation of the large and small domains. The only segment that moves significantly is a helix and loop that extends from residue Gly^{10} to Gly^{26} and moves by ∼2 Å where this movement appears to be a consequence of the induction of order in residues Leu^3−Ala^7. These are the first secondary structural elements of the small domain.

CobD is a dimer, as are all members of the aspartate aminotransferase family, where the active site lies close to the interface and is built from mostly one subunit, but with important contributions from the dimer-related protomer. In most aspartate aminotransferases, the N-terminal section of the polypeptide chain (usually α-helical) resides on the surface of the dimer-related subunit then extends across the dimer interface before it enters the small domain. In the process, it closes off the entrance to its active site. Typically, this section consists of >20 amino acid residues. In the structure of the PLP·CobD complex, the first seven amino acids were disordered, but even so, the missing residues were too few to form a major secondary structural element that might cross over to the second subunit. In the apo-CobD structure, five of the previously missing residues are visible, and surprisingly, they do not extend to the second subunit but instead fold back to lie against the small domain of the same subunit (Figure 3). In this way, the side chains of Leu^3 and Phe^4 are directed into a hydrophobic pocket at the base of which lie Leu^{316} and Leu^{321}. Thus, it would appear that the orientation of the N-terminal segment is not an artifact of crystallization.

Apart from the movement associated with the N-terminus, there are few other changes in the active site relative to the PLP·CobD complex. The phosphate ions present in the apo structure bind in essentially the same location as the phosphate in the PLP moiety and at the same site that was presumed to correspond to the substrate phosphate seen in the earlier structure (Figure 4A). The identity of this latter site was confirmed by the structure of the apo-CobD·L-threonine phosphate complex.

Structure of the Apo-CobD·L-Threonine Phosphate Complex. The polypeptide chain extends continuously from Phe^4 to Ala^{362} in the apo-CobD·substrate complex. This structure is also very similar to the PLP·CobD complex where the rms difference between 333 equivalent α-carbon atoms is 0.271 Å. The electron density for the substrate is unambiguous (Figure 2B). There are no major differences in the orientations of the side chains that form the active site, except that Lys^{216} (the residue that forms the internal aldimine to PLP) is partly disordered in the apo-CobD·substrate complex. As seen in Figure 4B, the substrate binds in the active site pocket with its phosphate moiety very close to that of the phosphate ion observed in the PLP·CobD structure. The coordination of the phosphate is essentially identical to that of the phosphate ion.

The ligand coordination diagram shown in Figure 5A

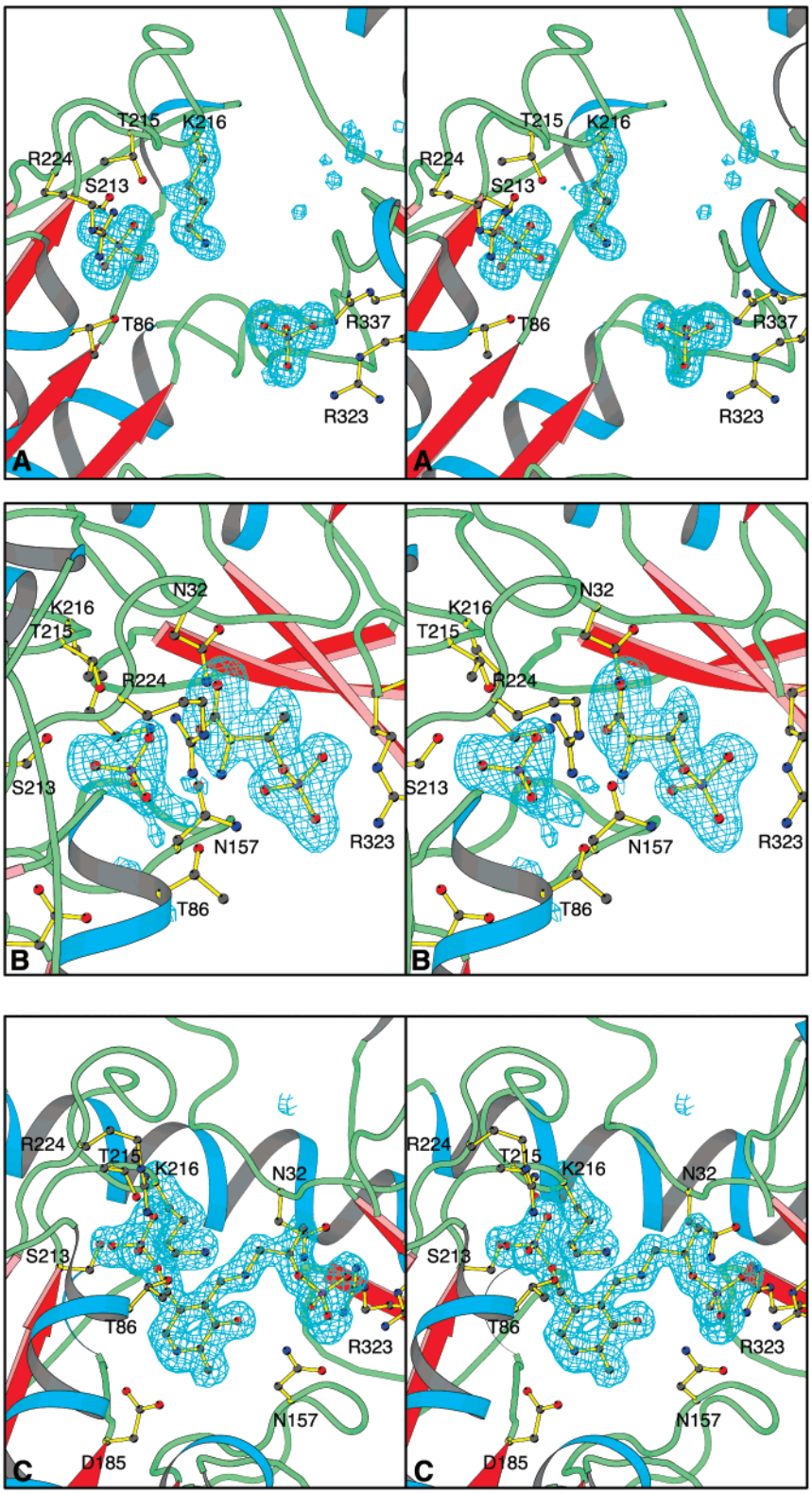


FIGURE 2: Stereoview of the electron density for the three structures of CobD: (A) a section of the electron density in apo-CobD for residues adjacent to the active site, (B) electron density for the substrate coordinated to the apoenzyme, and (C) electron density for the product intermediate. All maps were computed with coefficients of the form $F_o - F_c$ where the ligands and amino acid residues were excluded from the phase calculation and refinement. This figure was prepared with the programs Molscript (*38*) and Bobscript (*39*).

shows that most of the interactions with the substrate occur with the phosphate moiety that is highly coordinated. In contrast, there are only two interactions with the carboxylate group of L-threonine phosphate that occur between O1 and Nδ2 of Asn^{32} and O2 and the amide hydrogen of Asn^{32} with distances of 2.9 and 3.0 Å, respectively. Thus, it would appear that the position of the substrate in the active site is dominated by the phosphoryl interactions, which in turn control the relationship between the substrate and PLP. The absence of extensive hydrogen bonding interactions to the carboxylate probably favors the elimination of carbon dioxide.

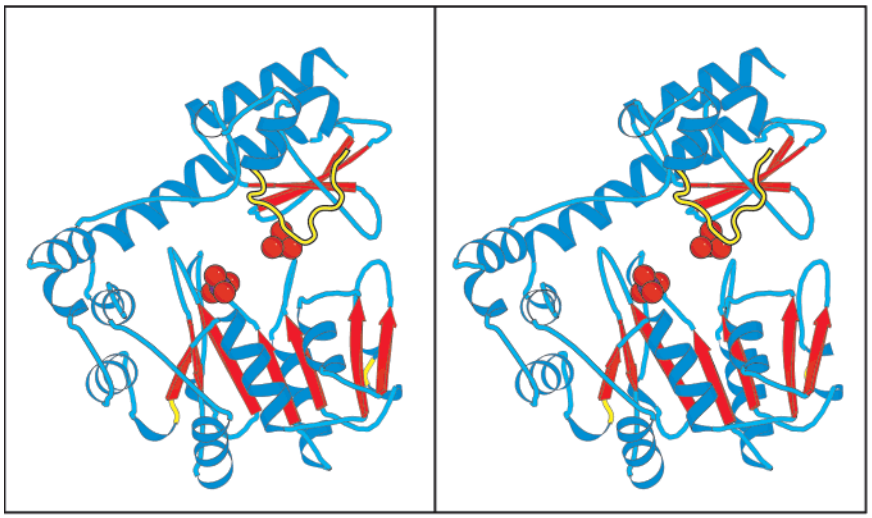


FIGURE 3: Stereoview of the apo-CobD monomer. The 12 N-terminal amino acids are colored in yellow to indicate that the polypeptide chain folds back on itself rather than extending to the dimer-related subunits as in other aspartate aminotransferases. The locations of the two phosphate ions in the active site are shown in a space filling representation. This figure was prepared with the program Molscript (*38*).

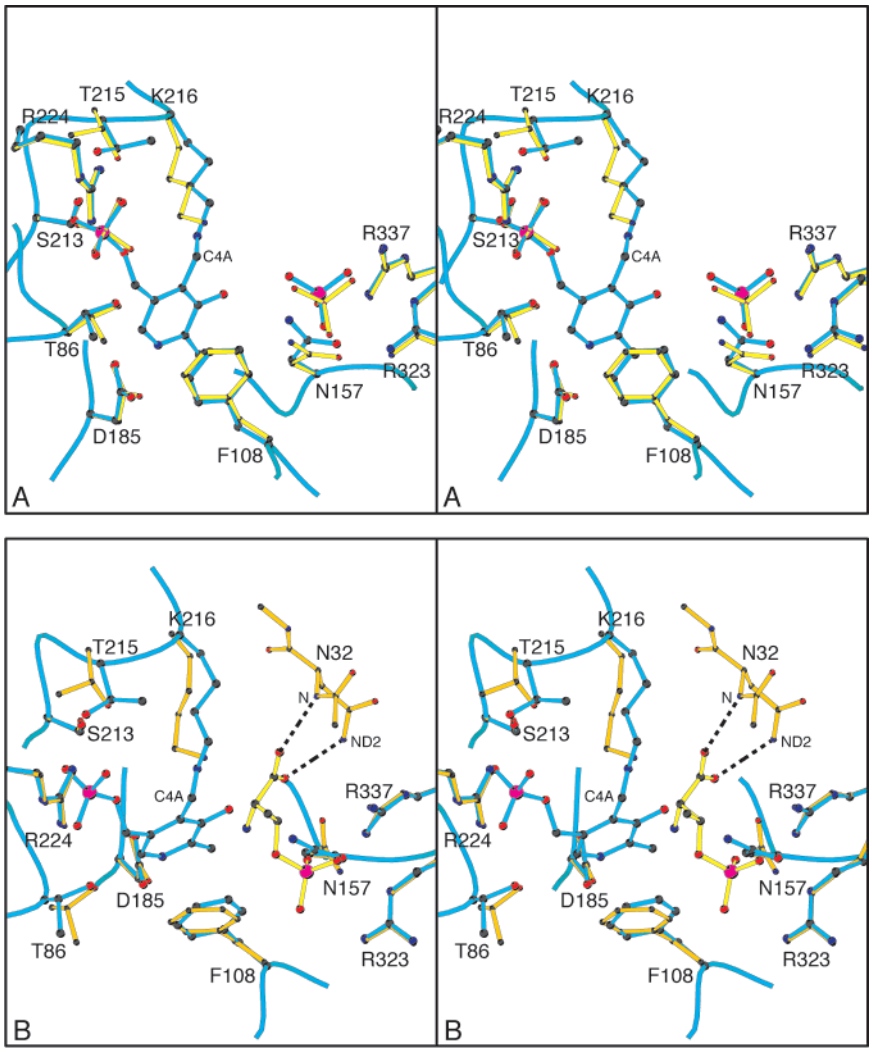


FIGURE 4: Stereo overlap of the wild-type CobD with (A) apo-CobD and (B) the apo-CobD·substrate complex. The structures were overlapped with the program Align (*40*) where the coordinates of the wild-type protein corresponded to PDB entry 1KUS (*8*). The wild-type protein is depicted with blue bonds and a blue connecting coil in both panels, whereas the side chains and contents of the active site for the apoprotein and apo-CobD·substrate complex are shown in yellow in panels A and B, respectively. This figure was prepared with the program Molscript (*38*).

The overlay of the apo-CobD·substrate complex and the PLP·CobD structure shows that the amino group of the L-threonine phosphate is close to an appropriate location for nucleophilic attack on C4′ of the internal aldimine (Figure

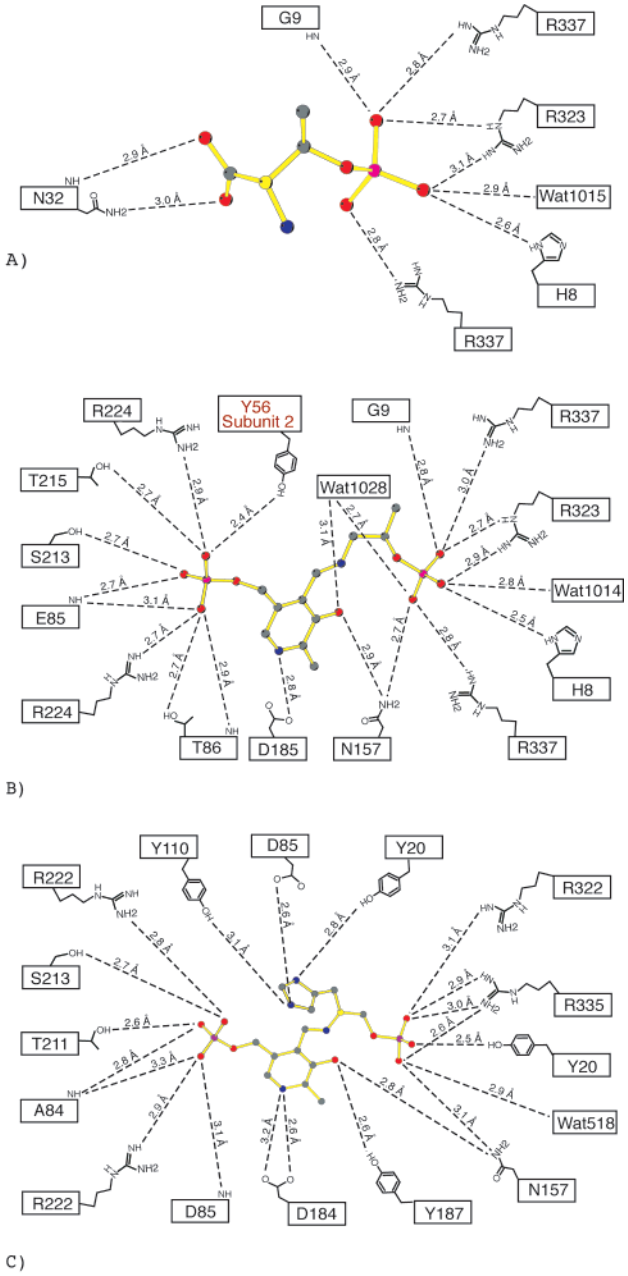


FIGURE 5: Ligand coordination diagrams of the binding interactions of (A) the apo-CobD·substrate complex, (B) the CobD·product aldimine complex, and (C) the ketimine complex between histidinol phosphate and PLP observed in histidinol phosphate aminotransferase. The coordinates for the ketimine complex of histidinol phosphate aminotransferase correspond to PDB entry 1GEX (*25*).

4B). In addition, the carboxylate moiety of the substrate is directed toward the general location of Lys^{216}.

According to the accepted mechanism for PLP-dependent enzymes, the fate of the external aldimine is controlled by the orientation of the substituents at the α-carbon relative to the plane of the coenzyme−imine π system. The bond that is broken is suggested to lie orthogonal to this plane so its electrons can be delocalized into the π system (*13*, *21*). Thus, in PLP-dependent enzymes, the reaction path is controlled by stereochemical steering. The orientation of the carboxylate group of the substrate is consistent with this theory since it is directed away from the plane of the PLP in part through its interactions with Asn^{32}. Examination of the active site reveals that once the position and orientation of the phosphate of the substrate have been established there is essentially nowhere else to place the remainder of the substrate and still establish an interaction between its amino group and the C4′ of PLP. Any other rotational orientation places the carboxylate of the substrate in a highly hydrophobic environment which would destabilize substrate binding or in collision with hydrophobic residues such as Phe^{108} or $Tyr^{56'}$ and $Trp^{246'}$ from the symmetry-related subunit.

An important consequence of the observed orientation for the carboxylate of the substrate is that it places the hydrogen on the α-carbon of L-threonine phosphate on the side of the pyridoxal ring opposite Lys^{216}. This orientation also contributes to selection of the decarboxylation reaction over deprotonation because the lysine that forms the internal aldimine is also believed to function as the base for proton abstraction in aspartate aminotransferases (*22*). In this observed orientation, the hydrogen cannot be abstracted by Lys^{216} and there are no other amino acid residues that can adopt that function within a 5 Å radius. Thus, the active site favors decarboxylation over hydrogen abstraction by providing an appropriate environment for the carboxylate, sterically inhibiting any other orientation and exclusion of the base for proton abstraction.

Structure of the Product Aldimine Complex. The electron density extends continuously from Ala^{7} to Pro^{362} in the product intermediate complex. Again, the overall fold of this complex is very similar to that of the PLP·CobD structure where the rms difference between 336 equivalent α-carbon atoms is 0.275 Å. In combination with the structure of apo-CobD and the substrate complex, this shows that the enzyme does not undergo a large conformational change when substrate binds and is thus similar to most other members of the α-family of aminotransferases (*23*−*25*), but is dissimilar to aspartate aminotransferase (*26*, *27*).

The electron density for the intermediate complex is unequivocal (Figure 2C). As can be seen, the nitrogen of the 1-amino-2-propanol phosphate lies virtually coplanar with the coenzyme ring and is positioned to form a hydrogen bond to O3′ of the pyridine ring. This geometry was not imposed in the stereochemical restraints of the refinement, but rather is a feature of the electron density. From this, it can be deduced that the species present in the active site is the external aldimine for the product.

Comparison of the product aldimine with PLP of the wild-type enzyme shows that the coenzyme tilts by approximately 13° on formation of the product (Figure 6A). This type of reorganization has been seen in other members of the α-family of PLP-dependent enzymes such as ornithine aminotransferase dialkyglycine decarboxylase, where it is suggested to arise from the need to avoid ionic and steric interaction with the lysine that forms the internal aldimine (*23*, *24*, *28*).

The interactions of the product aldimine with the active site are very similar to a combination of those for PLP·CobD and the CobD·substrate complex (Figure 5B). As expected, the interactions are dominated by ionic and hydrogen bonding contacts with the phosphate moieties of the PLP and 1-amino-2-propanol phosphate. There are no additional polar interactions with the amino-2-propanol moiety. Indeed, apart from interactions with the phosphate moiety, there are very few contacts with the aminopropanol group. The only nonpolar interaction is between the methyl group of the

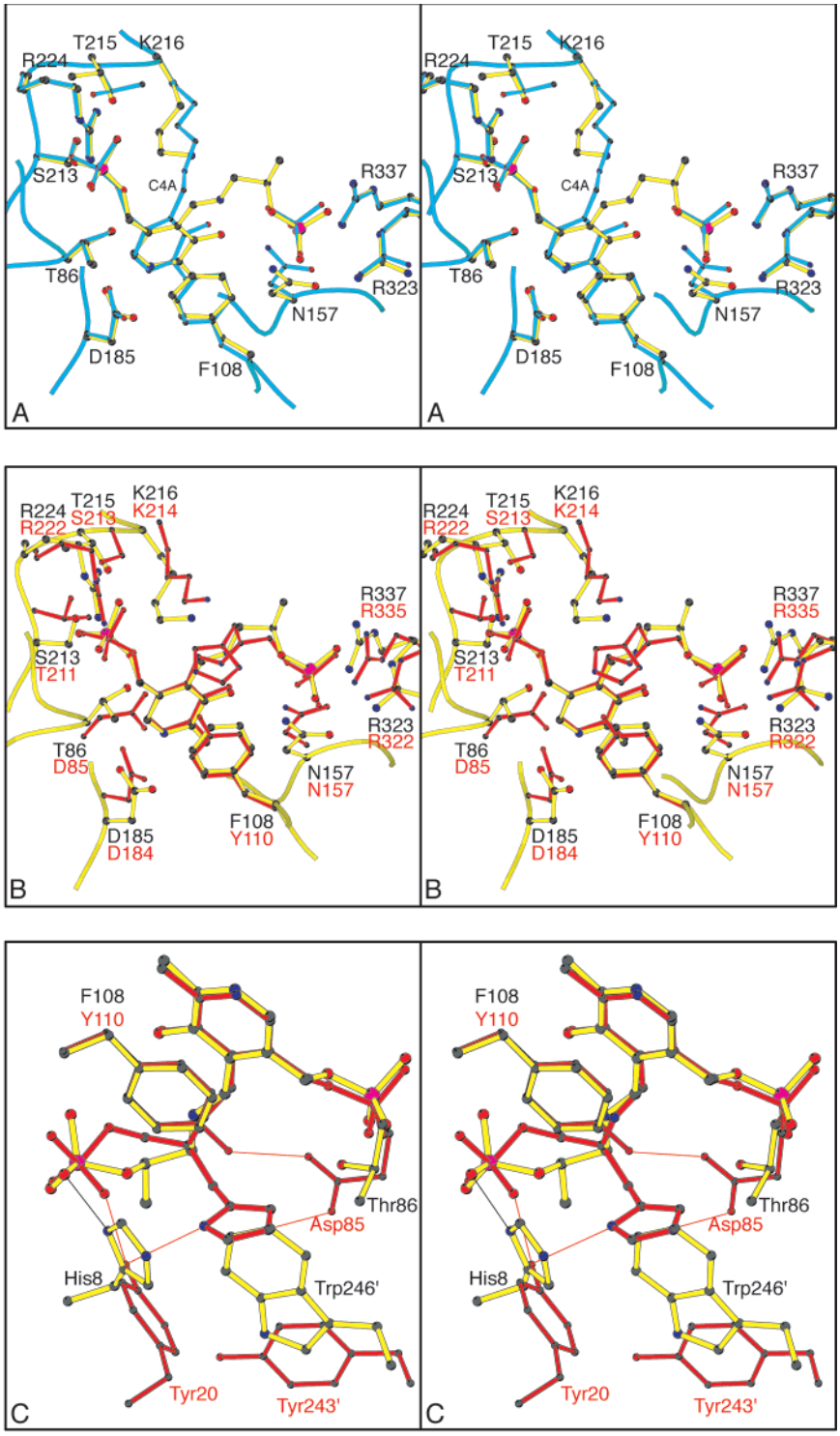


FIGURE 6: Stereo overlap of the product aldimine complex with (A) wild-type CobD and (B and C) histidinol phosphate aminotransferase. The structures were overlapped with the program Align (*40*) where the coordinates of the histidinol phosphate aminotransferase ketimine complex corresponded to PDB entry 1GEX (*25*). In panel A, the wild-type cobD·PLP complex is depicted with blue bonds and a blue coil, whereas the side chains and contents of the active site for the product aldimine complex are colored in yellow. In panel B, the product aldimine complex is colored in yellow whereas the ketimine intermediate for histidinol phosphate aminotransferase is depicted in red. For clarity, the side chains from the neighboring subunit and the N-terminal residues have been omitted. These latter residues are depicted in panel C, which shows those side chains that serve to discriminate between L-threonine phosphate and histidinol phosphate. This figure was prepared with the program Molscript (*38*). The coordinates for the ketimine complex of histidinol phosphate aminotransferase correspond to PDB entry 1GEX (*25*).

product and the α-carbon of Gly^{9}. Significantly, Lys^{216} lies 3.0 Å from the nitrogen atom of the nascent 1-amino-2-propanol phosphate. By analogy to the mechanism proposed for the aspartate aminotransferases, Lys^{216} is expected to serve as the general acid in the resolution of the product complex (*22*).

Comparison of the Active Sites of CobD with Histidinol Phosphate Aminotransferase. As noted earlier (*8*), the closest

structural and sequence relative to CobD is histidinol phosphate aminotransferase. This latter enzyme catalyzes the transfer of an amino group from histidinol phosphate to 2-oxoglutarate to form imidazole acetol phosphate and glutamate. Thus, in contrast to CobD, this enzyme must bind two substrates. Three structures of histidinol phosphate aminotransferase from *Escherichia coli* have been published: the wild-type enzyme and complexes with histidinol phosphate and *N*-(5′-phosphopyridoxyl)-L-glutamate (*25*). These provide an explanation for how both substrates can be accommodated in the active site in a manner that favors abstraction of a proton from the external aldimine rather than steering the reaction toward decarboxylation. These strategies are common to the entire aspartate aminotransferase family (*22*).

Strikingly, the coordination of the PLP moiety and phosphate of the product is essentially identical in both proteins. The differences arise in those side chains that accommodate the imidazole group of histidinol relative to the methyl group of 1-amino-2-propanol phosphate. The groups that differ significantly in the manner in which they interact with their ligands are His^{8}, Thr^{86}, and Trp^{246} in CobD which are spatially equivalent to Tyr^{20}, Asp^{85}, and Tyr^{243}, respectively, in histidinol phosphate aminotransferase. Of these, His^{8} in CobD interacts with a terminal phosphate oxygen (distance of 2.5 Å), whereas in histidinol phosphate aminotransferase, Tyr^{20} assumes this task, but in addition forms a second interaction with Nδ1 of the imidazole moiety of the ketimine intermediate (Figure 6). In a similar way, the side chain of Asp^{85} of histidinol phosphate aminotransferase forms a hydrogen bond to Nϵ2 of the imidazole moiety, whereas the equivalent side chain of CobD (Thr^{86}) is simply solvent exposed. An important difference between the enzymes is the substitution of $Tyr^{243'}$ in histidinol phosphate aminotransferase with $Trp^{246'}$ in CobD. This replacement serves to fill the space occupied by the imidazole moiety of the substrate in histidinol phosphate aminotransferase.

The strong similarity between CobD and histidinol phosphate aminotransferase suggests that it should be possible to convert each enzyme into its counterpart with only a modest number of changes (Figures 5C and 6C). In the first instance, replacement of the three side chains listed above with their equivalent residues in the target might be expected to remove the major impediments to accepting the alternative substrates; however, as seen in early attempts to alter specificity by site-directed mutagenesis of trypsin (*29*), it is difficult to predict if the new side chains will adopt a productive conformation. Thus, second-sphere substitutions are expected to be necessary to accommodate the new side chains. For example, Phe^{108} in CobD is equivalent to Tyr^{110} in histidinol phosphate aminotransferase, where the latter side chain forms a hydrogen bond between its hydroxyl group and the side chain of Asp^{85}. This secondary interaction serves to orient the side chain of Asp^{85} and thus contributes to the integrity of the binding site for the imidazole moiety of the substrate. Also, replacement of His^{8} with a tyrosine will not generate the functionality of Tyr^{20} in histidinol phosphate aminotransferase since the α-carbons of these side chains are in different locations in their respective structures (Figure 6C). In this case, examination of a number of hydrogen bonding side chains such as serine, threonine, or asparagine might be needed to determine whether it is possible to fulfill the role of the tyrosine with a simple amino acid replacement.

The residues listed above may also serve to indicate the identities of L-threonine decarboxylases and histidinol phosphate aminotransferases within the sequence databases. A BLAST search with the sequence for CobD as a probe reveals a substantial number of entries that have been annotated as histidinol phosphate aminotransferases and yet contain the residues equivalent to His^{8}, Thr^{85}, Phe^{108}, and Trp^{246} in CobD. For example, gi:11499606 in *Fusobacterium nucleatum* is annotated as a histidinol phosphate aminotransferase and yet contains all of the residues characteristic of CobD. Given the comparison of the structures of CobD and histidinol phosphate aminotransferase described here, it would appear to be unlikely that this protein is an aminotransferase considering also that the sequence surrounding these sites is also similar to that of CobD which implies they should have the same local structure.

Comparison of CobD with Other PLP-Dependent Decarboxylases. There are a large number of PLP-dependent decarboxylases in the α-family, though the sequences for most are distantly related to that of CobD relative to the aminotransferases. The closest decarboxylases, as determined by a search of the protein database (*30*, *31*) with the program DALI (*32*), are L-3,4-dihydroxyphenylalanine decarboxylase (DOPA decarboxylase) (*33*) and 2,2-dialkylglycine decarboxylase (*28*). These two enzymes exhibit *Z* scores of 25.2 and 17.8, respectively, whereas the corresponding value for histidinol phosphate aminotransferase was 36.4 (PDB entry 1GEW) (*25*).

Considerable effort has been devoted to understanding the structure and function of 2,2-dialkylglycine decarboxylase (*12*, *28*, *34*, *35*). This enzyme catalyzes oxidative decarboxylation of dialkylglycines such as 2-aminobutyrate to yield acetone and CO_2 followed by a conventional transamination of pyruvate to obtain L-alanine. Again, in contrast to CobD, this enzyme must recognize two substrates.

As expected, the overall structure of CobD is very similar to that of dialkylglycine decarboxylase. In particular, most of the features that coordinate the PLP moiety are very similar in both enzymes. The differences arise in the components of the active site that coordinate the substrates. Comparison of the active sites of CobD and dialkylglycine decarboxylase bound to a series of aldimine complexes and analogues (data not shown) reveals that the carboxylate of pyruvate occupies a location similar to that of the phosphate moiety in 1-amino-2-propanol phosphate (*28*). In this orientation, the relationship between PLP, pyruvate, and the essential lysine stereochemically restricts the final product to L-alanine.

A consequence of the necessity to bind two substrates in the same active site creates an unusual situation where the dialkylglycine can bind to the enzyme in three orientations (*36*), of which only one leads to decarboxylation. Experimental confirmation of the carboxylate binding site in DGD has thus far not been achieved; however, the restraints imposed by the stereochemistry of the substrate (*36*) place that group in a location analogous to that seen in CobD.

At this time, the only other decarboxylases whose structures are known and belong to the fold 1 family (α-family) are bacterial ornithine decarboxylase and L-3,4-dihydroxyphenylalanine decarboxylase (DOPA decarboxylase) (*33*, *37*).

Of these, the substrate for ornithine decarboxylase is structurally closer to L-threonine phosphate than L-3,4-dihydroxyphenylalanine is; however, no analogue complexes that contain a carboxylate moiety are available for the former ornithine decarboxylase. In the case of DOPA decarboxylase, a complex with the inhibitor carbiDOPA has been determined which demonstrates that the carboxylate lies out of the plane of the pyridoxal cofactor, but on the opposite side of the ring relative to that observed in the apo-CobD·substrate complex (ref *33* and data not shown). These studies confirm that the PLP-binding motif found in the α-family can facilitate decarboxylation reactions in which the carboxylate moiety is eliminated on either side of the pyridoxal ring.

CONCLUSIONS

The studies described here demonstrate that L-threonine-*O*-3-phosphate decarboxylase (CobD) from *S. enterica* is closely related to histidinol phosphate aminotransferase in the manner in which it binds substrate, but steers the breakdown of its external aldimine toward decarboxylation instead of amino transfer. This is accomplished by positioning the carboxylate moiety of L-threonine *O*-3-phosphate out of the plane of the pyridoxal ring and by placing the α-hydrogen out of the reach of the catalytic base provided by the lysine that forms the internal aldimine. It would appear that CobD evolved from a histidinol phosphate aminotransferase (or a more primordial PLP-dependent aminotransferase), where the selection was based on similarities between the stereochemical properties of the substrates rather than preservation of the fate of the external aldimine.

The remarkable similarity between CobD and histidinol phosphate aminotransferase raises the question of whether CobD is truly an ancient enzyme or a more recent newcomer based on gene duplication and mutagenesis of an aminotransferase. If this were true, then the earlier enzyme has been lost or superseded in *S. enterica*, but might still exist in other organisms. The argument against the recent evolution of CobD is the observation of genes in archaea that contain the residues that would appear to discriminate between a L-threonine-*O*-3-phosphate decarboxylase and a histidinol phosphate aminotransferase.

The similarity between CobD and histidinol phosphate aminotransferase suggests that it should be possible to interconvert the activities. Two approaches to this problem are possible. In the first instance, targeted mutagenesis of the dissimilar residues should reveal if our understanding of the substrate specificity is sufficient to correctly redesign each of the enzymes. In the second, the availability of bacterial strains that require these genes for survival should allow the use of random mutagenesis and selection for function in yielding a biological route for interconverting these biological activities. Efforts to examine the feasibility of both strategies are in progress.

REFERENCES

1. Rondon, M. R., Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1997) Biochemistry and Molecular Genetics of Cobalamin Biosynthesis, *Prog. Nucleic Acid Res. Mol. Biol. 56*, 347−384.
2. Scott, A. I. (2001) Reflections on the discovery of nature's pathways to vitamin B_{12}, *Chem. Rec. 1*, 212−227.
3. Story, R. M., Weber, I. T., and Steitz, T. A. (1992) The structure of the *E. coli* recA protein monomer and polymer, *Nature 355*, 318−325.
4. Abrahams, J. P., Leslie, A. G., Lutter, R., and Walker, J. E. (1994) Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria, *Nature 370*, 621−628.
5. Sawaya, M. R., Guo, S., Tabor, S., Richardson, C. C., and Ellenberger, T. (1999) Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7, *Cell 99*, 167−177.
6. Brushaber, K. R., O'Toole, G. A., and Escalante-Semerena, J. C. (1998) CobD, a novel enzyme with L-threonine-*O*-3-phosphate decarboxylase activity, is responsible for the synthesis of (*R*)-1-amino-2-propanol *O*-2-phosphate, a proposed new intermediate in cobalamin biosynthesis in *Salmonella typhimurium* LT2, *J. Biol. Chem. 273*, 2684−2691.
7. Mehta, P. K., and Christen, P. (1993) Homology of pyridoxal-5′-phosphate-dependent aminotransferases with the cobC (cobalamin synthesis), nifS (nitrogen fixation), pabC (p-aminobenzoate synthesis) and malY (abolishing endogenous induction of the maltose system) gene products, *Eur. J. Biochem. 211*, 373−376.
8. Cheong, C. G., Bauer, C. B., Brushaber, K. R., Escalante-Semerena, J. C., and Rayment, I. (2002) The Three-Dimensional Structure of the L-Threonine-O-3-phosphate Decarboxylase (CobD) Enzyme from *Salmonella enterica*, *Biochemistry 41*, 4798−4808.
9. Jansonius, J. N. (1998) Structure, evolution and action of vitamin B6-dependent enzymes, *Curr. Opin. Struct. Biol. 8*, 759−769.
10. Mehta, P. K., and Christen, P. (2000) The molecular evolution of pyridoxal-5′-phosphate-dependent enzymes, *Adv. Enzymol. Relat. Areas Mol. Biol. 74*, 129−184.
11. Capitani, G., Hohenester, E., Feng, L., Storici, P., Kirsch, J. F., and Jansonius, J. N. (1999) Structure of 1-aminocyclopropane-1-carboxylate synthase, a key enzyme in the biosynthesis of the plant hormone ethylene, *J. Mol. Biol. 294*, 745−756.
12. Toney, M. D., Hohenester, E., Cowan, S. W., and Jansonius, J. N. (1993) Dialkylglycine decarboxylase structure: bifunctional active site and alkali metal sites, *Science 261*, 756−759.
13. Dunathan, H. C. (1971) Stereochemical aspects of pyridoxal phosphate catalysis, *Adv. Enzymol. Relat. Areas Mol. Biol. 35*, 79−134.
14. Metzler, C. M. (2001) in *Biochemistry, The Chemical Reactions of Living Cells*, pp 740−752, Harcourt/Academic Press, San Diego.
15. Rayment, I. (2002) Small-scale batch crystallization of proteins revisited. An underutilized way to grow large protein crystals, *Structure 10*, 147−151.
16. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
17. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
18. Roussel, A., and Cambillau, C. (1991) in *Silicon Graphics Geometry Partners Directory*, Silicon Graphics Mountain View, CA.
19. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
20. Ford, G. C., Eichele, G., and Jansonius, J. N. (1980) Three-dimensional structure of a pyridoxal-phosphate-dependent enzyme, mitochondrial aspartate aminotransferase, *Proc. Natl. Acad. Sci. U.S.A. 77*, 2559−2563.
21. Dunathan, H. C. (1966) Conformation and reaction specificity in pyridoxal phosphate enzymes, *Proc. Natl. Acad. Sci. U.S.A. 55*, 712−716.
22. Kirsch, J. F., Eichele, G., Ford, G. C., Vincent, M. G., Jansonius, J. N., Gehring, H., and Christen, P. (1984) Mechanism of action of aspartate aminotransferase proposed on the basis of its spatial structure, *J. Mol. Biol. 174*, 497−525.
23. Shah, S. A., Shen, B. W., and Brunger, A. T. (1997) Human ornithine aminotransferase complexed with L-canaline and gabaculine: structural basis for substrate recognition, *Structure 5*, 1067−1075.
24. Storici, P., Capitani, G., Muller, R., Schirmer, T., and Jansonius, J. N. (1999) Crystal structure of human ornithine aminotransferase complexed with the highly specific and potent inhibitor 5-fluoromethylornithine, *J. Mol. Biol. 285*, 297−309.
25. Haruyama, K., Nakai, T., Miyahara, I., Hirotsu, K., Mizuguchi, H., Hayashi, H., and Kagamiyama, H. (2001) Structures of

Escherichia coli histidinol-phosphate aminotransferase and its complexes with histidinol-phosphate and N-(5′-phosphopyridoxyl)-L-glutamate: double substrate recognition of the enzyme, *Biochemistry 40*, 4633−4644.
26. McPhalen, C. A., Vincent, M. G., Picot, D., Jansonius, J. N., Lesk, A. M., and Chothia, C. (1992) Domain closure in mitochondrial aspartate aminotransferase, *J. Mol. Biol. 227*, 197−213.
27. Rhee, S., Silva, M. M., Hyde, C. C., Rogers, P. H., Metzler, C. M., Metzler, D. E., and Arnone, A. (1997) Refinement and comparisons of the crystal structures of pig cytosolic aspartate aminotransferase and its complex with 2-methylaspartate, *J. Biol. Chem. 272*, 17293−17302.
28. Malashkevich, V. N., Strop, P., Keller, J. W., Jansonius, J. N., and Toney, M. D. (1999) Crystal structures of dialkylglycine decarboxylase inhibitor complexes, *J. Mol. Biol. 294*, 193−200.
29. Graf, L., Craik, C. S., Patthy, A., Roczniak, S., Fletterick, R. J., and Rutter, W. J. (1987) Selective alteration of substrate specificity by replacement of aspartic acid-189 with lysine in the binding pocket of trypsin, *Biochemistry 26*, 2616−2623.
30. Berman, H. M., Bhat, T. N., Bourne, P. E., Feng, Z., Gilliland, G., Weissig, H., and Westbrook, J. (2000) The Protein Data Bank and the challenge of structural genomics, *Nat. Struct. Biol. 7*, 957−959.
31. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures, *J. Mol. Biol. 112*, 535−542.
32. Holm, L., and Sander, C. (1999) Protein folds and families: sequence and structure alignments, *Nucleic Acids Res. 27*, 244−247.
33. Burkhard, P., Dominici, P., Borri-Voltattorni, C., Jansonius, J. N., and Malashkevich, V. N. (2001) Structural insight into Parkinson's disease treatment from drug-inhibited DOPA decarboxylase, *Nat. Struct. Biol. 8*, 963−967.
34. Hohenester, E., Keller, J. W., and Jansonius, J. N. (1994) An alkali metal ion size-dependent switch in the active site structure of dialkylglycine decarboxylase, *Biochemistry 33*, 13561−13570.
35. Zhou, X., Kay, S., and Toney, M. D. (1998) Coexisting kinetically distinguishable forms of dialkylglycine decarboxylase engendered by alkali metal ions, *Biochemistry 37*, 5761−5769.
36. Toney, M. D., Hohenester, E., Keller, J. W., and Jansonius, J. N. (1995) Structural and mechanistic analysis of two refined crystal structures of the pyridoxal phosphate-dependent enzyme dialkylglycine decarboxylase, *J. Mol. Biol. 245*, 151−179.
37. Momany, C., Ernst, S., Ghosh, R., Chang, N. L., and Hackert, M. L. (1995) Crystallographic structure of a PLP-dependent ornithine decarboxylase from *Lactobacillus* 30a to 3.0 Å resolution, *J. Mol. Biol. 252*, 643−655.
38. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
39. Esnouf, R. M. (1999) Further additions to MolScript version 1.4, including reading and contouring of electron-density maps, *Acta Crystallogr. D55*, 938−940.
40. Cohen, G. H. (1997) ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions, *J. Appl. Crystallogr. 30*, 1160−1161.

BI020294W