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-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-3-phosphate decarboxylase and histidinol phosphate aminotransferases, many of which appear to have been misannotated.
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 aminotransferase 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 concentrated to 10 mg/mL. A sodium borohydride solution (100 mM NaCl) was prepared at 4 °C for future use.

**Preparation of the Apoenzyme.** Efforts to inactivate the enzyme by reducing the Schiff base between PLP and Lys216 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 °C and frozen dropwise in liquid nitrogen and stored at -80 °C for future use.

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

![Diagram](image-url)
Complexes of L-Threonine-O-3-phosphate Decarboxylase

°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 spectrophotometrically 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 I222 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 CobD Reaction Intermediate Complex.** Crystals of apo-CobD in complex with l-threonine phosphate 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 mM 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 I222 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 (Rfree). Water molecules were incorporated into the structure if they had peak heights of >3σ in the Fo − Fc 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 1: Data Collection Statistics**

<table>
<thead>
<tr>
<th></th>
<th>apo-CobD</th>
<th>substrate complex</th>
<th>reaction intermediate complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>resolution (Å)</td>
<td>1.46</td>
<td>1.8</td>
<td>1.8</td>
</tr>
<tr>
<td>no. of unique reflections</td>
<td>74541</td>
<td>38108</td>
<td>37471</td>
</tr>
<tr>
<td>redundancy</td>
<td>8.2</td>
<td>8.8</td>
<td>7.6</td>
</tr>
<tr>
<td>completeness (%)</td>
<td>99.4 (95.4)</td>
<td>99.6 (99.1)</td>
<td>99.4 (99.2)</td>
</tr>
<tr>
<td>average</td>
<td>66.7 (13.2)</td>
<td>60.8 (10.2)</td>
<td>31.9 (5.8)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.040 (0.129)</td>
<td>0.055 (0.219)</td>
<td>0.078 (0.331)</td>
</tr>
</tbody>
</table>

* The numbers in parentheses represent completeness in the highest-resolution shell. * Rmerge = ∑(i)(j) ∑(j) I(j)−I(i) ∑(i)I(j), where the average intensity I is taken over all symmetry equivalent measurements and I(i) is the measured intensity for a given reflection.
Table 2: Refinement Statistics

<table>
<thead>
<tr>
<th></th>
<th>apo-CobD substrate complex</th>
<th>product aldmine complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>resolution limits (Å)</td>
<td>1.46 1.8</td>
<td>1.8</td>
</tr>
<tr>
<td>R-factor</td>
<td>19.0 19.6</td>
<td>19.6</td>
</tr>
<tr>
<td>R_rms</td>
<td>21.4 21.2 22.9</td>
<td></td>
</tr>
<tr>
<td>no. of reflections</td>
<td>70826 35941 35608</td>
<td></td>
</tr>
<tr>
<td>no. of reflections</td>
<td>3715 1908 1862</td>
<td></td>
</tr>
<tr>
<td>no. of protein atoms</td>
<td>2775 2733 2735</td>
<td></td>
</tr>
<tr>
<td>no. of solvent atoms</td>
<td>271 228</td>
<td></td>
</tr>
<tr>
<td>no. of molecules, ions</td>
<td>2 PO4</td>
<td></td>
</tr>
<tr>
<td>electron density</td>
<td></td>
<td></td>
</tr>
<tr>
<td>average B value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>main chain atoms</td>
<td>14.4 25.8 19.7</td>
<td></td>
</tr>
<tr>
<td>all protein atoms</td>
<td>16.4 27.4 21.4</td>
<td></td>
</tr>
<tr>
<td>solvent atoms</td>
<td>25.4 33.6 28.5</td>
<td></td>
</tr>
<tr>
<td>ligand atoms</td>
<td>10.4 30.4 14.6</td>
<td></td>
</tr>
<tr>
<td>weighted rms deviations from ideality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond lengths (Å)</td>
<td>0.011 0.011 0.013</td>
<td></td>
</tr>
<tr>
<td>bond angles (deg)</td>
<td>1.54 1.53 1.57</td>
<td></td>
</tr>
<tr>
<td>PDB entry</td>
<td>1LC5 1LC7 LLC8</td>
<td></td>
</tr>
</tbody>
</table>

* CNS refinement. \( R = \frac{\sum |F_{o} - \bar{F}_{o}| \sum |\bar{F}_{o}|}{|F_{o}|} \), where 5% of the data were excluded for calculation of \( R_{\text{free}} \).

and an \( R_{\text{free}} \) 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 aldmine 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 (\( R_{\text{free}} \)). Water molecules were incorporated into the structure if they had peak heights of > 3\( \sigma \) in the \( F_{o} - \bar{F}_{o} \) 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 \( R_{\text{free}} \) values of 25.1 and 22.9% for substrate and product aldmine 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 aldmine 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 \( \alpha \)-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\(^8\). 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 \( \alpha \)-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\(^8\) and Phe\(^4\) are directed into a hydrophobic pocket at the base of which lie Leu\(^{136}\) and Leu\(^{251}\). 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\(_{\text{L-threonine phosphate complex}}\).

Structure of the Apo-CobD\(_{\text{L-threonine phosphate complex}}\). The polypeptide chain extends continuously from Phe\(^4\) to Ala\(^{362}\) in the apo-CobD\(_{\text{substrate complex}}\). This structure is also very similar to the PLP-CobD complex where the rms difference between 333 equivalent \( \alpha \)-carbon atoms is 0.271 Å. The electron density for the substrate is unambiguously (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 aldmine to PLP) is partly disordered in the apo-CobD\(_{\text{substrate complex}}\). 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.
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 N2 of Asn32 and O2 and the amide hydrogen of Asn32 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.
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 4).
The interactions of the product aldimine with the active site are very similar to a combination of those for PLP-CobD and the 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 contacts with the phosphate moiety of the PLP and the CobD 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 for the product.

Structure of the Product Aldimine Complex. The electron density extends continuously from Ala7 to Pro362 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 stereoechemical 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 2A). 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 substrate and the general location of Lys216.

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 Asn32. 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 Phe108 or Tyr56 and Trp246 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 Lys216. 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 Lys216 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.

Figure 5: Ligand coordination diagrams of the binding interactions of (A) the apo-CobD-substrate complex, (B) the CobD-product aldime complex, and (C) the ketimine complex between histidinol phosphate aminotransferase. The coordinates for the ketimine complex of histidinol phosphate aminotransferase correspond to PDB entry 1GEX (25).
product and the α-carbon of Gly. Significantly, Lys 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 is expected to serve as the general acid in the resolution of the product complex.

Comparison of the Active Sites of CobD with Histidinol Phosphate Aminotransferase. As noted earlier, 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 acetyl 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′-phosphorpyridoxyl)-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 His8, Thr86, and Trp246 in CobD which are spatially equivalent to Tyr20, Asp85, and Tyr243, respectively, in histidinol phosphate aminotransferase. Of these, His8 in CobD interacts with a terminal phosphate oxygen (distance of 2.5 Å), whereas in histidinol phosphate aminotransferase, Tyr20 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 Asp85 of histidinol phosphate aminotransferase forms a hydrogen bond to Ne2 of the imidazole moiety, whereas the equivalent side chain of CobD (Thr86) is simply solvent exposed. An important difference between the enzymes is the substitution of Tyr243 in histidinol phosphate aminotransferase with Trp246 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, Phe108 in CobD is equivalent to Tyr110 in histidinol phosphate aminotransferase, where the latter side chain forms a hydrogen bond between its hydroxyl group and the side chain of Asp85. This secondary interaction serves to orient the side chain of Asp85 and thus contributes to the integrity of the binding site for the imidazole moiety of the substrate. Also, replacement of His8 with a tyrosine will not generate the functionality of Tyr20 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 His8, Thr86, and Trp246 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 CO₂ 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 \( \alpha \)-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 \( \alpha \)-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**

Complexes of L-Threonine-O-3-phosphate Decarboxylase

Biochemistry, Vol. 41, No. 29, 2002  9089


BI020294W