The Three-Dimensional Structures of Nicotinate Mononucleotide:5,6-Dimethylbenzimidazole Phosphoribosyltransferase (CobT) from *Salmonella typhimurium* Complexed with 5,6-Dimethybenzimidazole and Its Reaction Products Determined to 1.9 Å Resolution[†,‡]

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

Institute for Enzyme Research and Department of Biochemistry, University of Wisconsin, Madison, Wisconsin

Received July 28, 1999; Revised Manuscript Received September 23, 1999

ABSTRACT: Nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase (CobT) from *Salmonella typhimurium* plays a central role in the synthesis of α-ribazole, which is a key component of the lower ligand of cobalamin. Two X-ray structures of CobT are reported here at 1.9 Å resolution. First, a complex of CobT with 5,6-dimethylbenzimidazole, and second, a complex of CobT with its reaction products, nicotinate and α-ribazole-5′-phosphate. CobT was cocrystallized with 5,6-dimethylbenzimidazole (DMB) in the space group $P2_12_12$ with unit cell dimensions of $a = 72.1$ Å, $b = 90.2$ Å, and $c = 47.5$ Å and one protomer per asymmetric unit. Subsequently, the crystals containing DMB were soaked in nicotinate mononucleotide whereupon the physiological reaction occurred in the crystal lattice to yield nicotinate and α-ribazole-5′-phosphate. These studies show that CobT is a dimer where each subunit consists of two domains. The large domain is dominated by a parallel six-stranded β-sheet with connecting α-helices that exhibit the topology of a Rossmann fold. The small domain is made from components of the N- and C-terminal sections of the polypeptide chain and contains a three-helix bundle. The fold of CobT is unrelated to the type I and II phosphoribosylpyrophosphate dependent transferases and does not appear to be related to any other protein whose structure is known. The enzyme active site is located in a large cavity formed by the loops at the C-terminal ends of the β-strands and the small domain of the neighboring subunit. DMB binds in a hydrophobic pocket created in part by the neighboring small domain. This is consistent with the broad specificity of this enzyme for aromatic substrates [Trzebiatowski, J. R., Escalante-Semerena (1997) *J. Biol. Chem. 272,* 17662−17667]. The binding site for DMB suggests that Glu^{317} is the catalytic base required for the reaction. The remainder of the cavity binds the nicotinate and ribose-5′-phosphate moieties, which are nestled within the loops at the ends of the β-strands. Interestingly, the orientation of the substrate and products are opposite from that expected for a Rossmann fold.

Cobalamin is the largest and most complex cofactor found in biological systems and has attracted considerable attention since the discovery that its deficiency is the cause of pernicious anemia (for a historical review, see 1). Cobalamin consists of a corrin ring that provides four of the ligands that coordinate a central cobalt ion (Figure 1). The two remaining coordination sites are usually defined as the upper and lower ligands. When the cofactor is in the active form, the upper ligand is either a 5′-deoxyadenosyl group or a methyl group, depending on its enzymatic role. The lower ligand is commonly dimethylbenzimidazole ribonucleotide, which is connected to the corrin ring by amino-2-propanol; however, there is considerable variation in the nature of the nucleotide base (*2*). The entire lower ligand assembly is often referred to as the nucleotide loop.

There have been two major milestones in understanding this magnificent molecule: first, the determination of its three-dimensional structure (*3*) and second, dissection of its biosynthetic pathway in *Pseudomonas denitrificans* (*4, 5*). Studies of the biosynthesis of cobalamin in a variety of organisms show that bacteria devote a significant amount of their genomic material to the synthesis of this complex molecule (*6*). This serves to underlie its biological and evolutionary significance. Even with this outstanding progress, there remain many important unanswered questions concerning the structure, function, and inter-relationship of the enzymes involved in cobalamin biosynthesis. For example, comparatively little is known of the three-dimensional structures responsible for biosynthesis of cobalamin and the precursors necessary for its assembly. This paper attempts to address this issue by describing the structure of nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltrans-

[⊥] Department of Bacteriology, University of Wisconsin, Madison, Wisconsin, USA.

* To whom correspondence may be addressed. Institute for Enzyme Research, 1710 University Avenue, Madison, WI 53705. Telephone: 608-262-0529. Fax: 608-262-1319. E-mail: ivan_rayment@biochem.wisc.edu.

[†] This research was supported in part by NIH grants AR35186 and GM58281 to I. R. and GM40313 to J.E−S.

[‡] The X-ray coordinates have been deposited in the Protein Data Bank with file names 1D0S and 1D0V for the complex of CobT with 5,6-Dimethybenzimidazole and the reaction products, respectively.

[1] Abbreviations: DTT, Dithiothreitol; EDTA, Ethylenediaminetetraacetic Acid; rms, root-mean-square; DMB, 5,6-dimethylbenzimidazole; FMN, flavin mononucleotide; NaMN, Nicotinate mononucleotide.

10.1021/bi991752c CCC: $18.00
Published on Web 11/16/1999

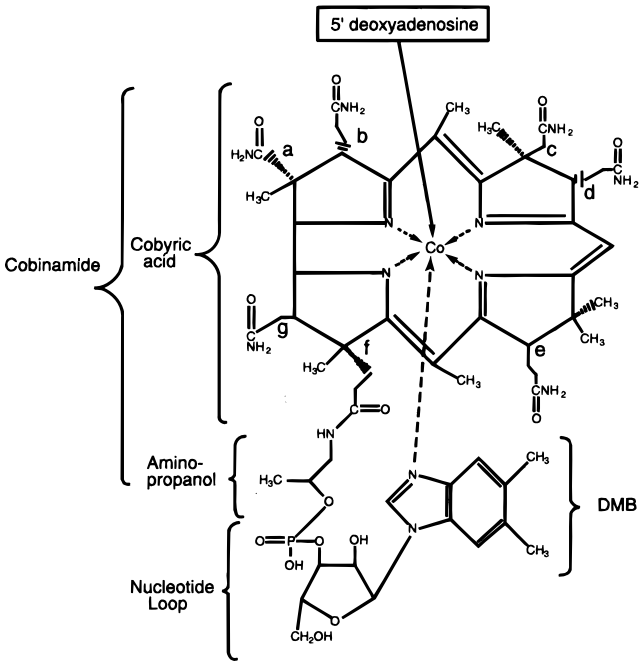


FIGURE 1: Chemical structure of adenosylcobalamin.

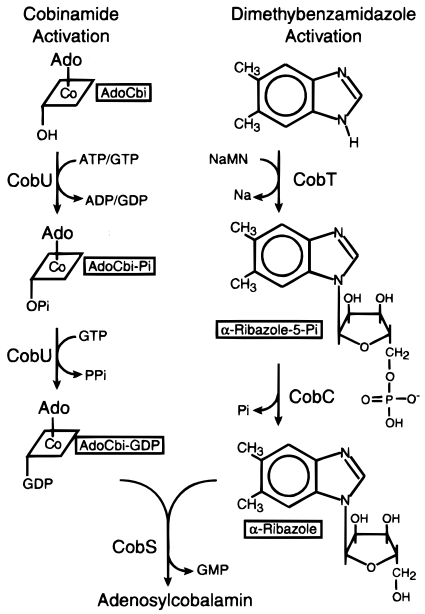
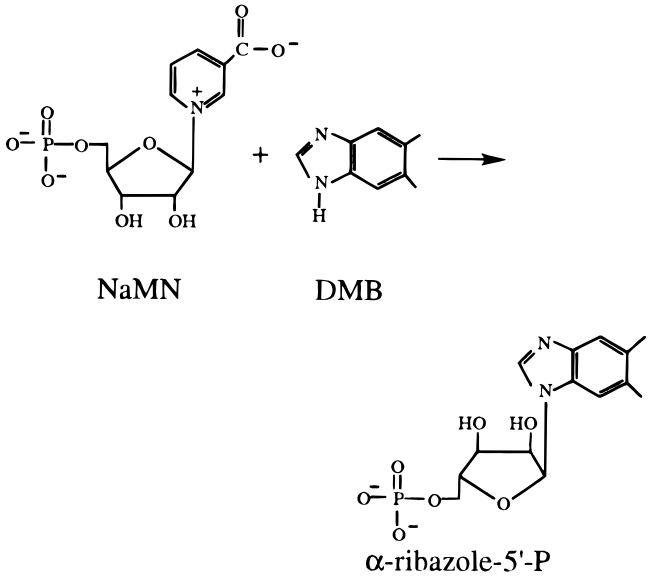


FIGURE 2: A schematic chemical representation of nucleotide loop synthesis and assembly in adenosylcobalamin biosynthesis in *Salmonella typhimurium*. Reproduced with permission from Thompson et al. (*53*).

ferase, the enzyme responsible for activating the lower ligand base.

Much of the current information regarding the biosynthesis of this complex cofactor has been derived from experiments with the bacterial systems, *P. denitrificans* and *Salmonella typhimurium* (*2, 4, 5, 7*). For convenience, the biosynthetic pathway can be broken down into the following: (i) synthesis of the corrin ring, (ii) attachment of the upper 5′-deoxyadenosine ligand to the cobalt ion, (iii) synthesis of the lower 5,6-dimethylbenzimidazole ligand, and (iv) assembly of the nucleotide loop. In *S. typhimurium*, there are four enzymes required to synthesize the nucleotide loop, and to assemble the lower ligand of cobalamin (Figure 2): CobU, CobT, CobC, and CobS (Figure 2). Of these, CobT and CobC are

Scheme 1



responsible for synthesizing the nucleoside of the lower ligand.

Initial synthesis and activation of the lower ligand (Figure 2) occurs through the transfer of the phosphoribosyl group of nicotinate mononucleotide to 5,6-dimethylbenzimidazole to yield α-ribazole by the action of CobT and CobC. The *cobT* gene of *S. typhimurium* has been demonstrated to code for a polypeptide chain of molecular weight 39 000 that contains 356 amino acid residues (*8, 9*). CobT catalyzes the synthesis of α-ribazole-5′-phosphate (Scheme 1), which is subsequently dephosphorylated by CobC (*10*). This reaction is unusual because it requires nicotinate mononucleotide in place of phosphoribosylpyrophosphate, which is used in most phosphoribosyltransferases (*11*). CobT shares no sequence similarity with any other phosphoribosyltransferase and is expected to have a different fold from any of the known structures (*12, 13*). The *cobC* gene from *S. typhimurium* has been cloned, sequenced, and shown to encode a polypeptide chain of molecular weight 34 000 that functions as a phosphatase (*10*). Interestingly, the amino acid sequence of CobC shows striking similarity to those of phosphoglycerate mutase and eukaryotic fructose kinase/bisphosphatase (*10, 14*−*16*). The final step in the nucleotide loop assembly pathway, the joining of adenosylcobinamide-GDP and α-ribazole, is catalyzed by CobS (*8, 9*). The *cobS* gene of *S. typhimurium* has been sequenced (*9*) and identified as a cobalamin synthase (*17*).

The role of CobT as a phosphoribosyltransferase in the biosynthesis of α-ribazole-5′-phosphate has been firmly established (*11*); however, additional functions have been ascribed to this enzyme. In a controversial proposal by Chen et al. (*18*), it has been suggested that CobT might be solely responsible for conversion of flavin mononucleotide (FMN) to DMB. There is good reason to believe that DMB is synthesized from FMN in the aerobic pathway (*19*−*21*). Indeed, recent work has shown that the C2 carbon of DMB is derived from the ribosyl moiety of FMN (*22*). The proposed pathway requires at least eight enzymatic steps. Mutations that render *S. typhimurium* unable to make adenosylcobalamin unless provided with DMB have been isolated and characterized genetically (*23*) and physically (*24*). Interest-

ingly, all of these mutations map to the *cobT* gene (*18*, *25*). Trzebiatowski et al. proposed two models to explain the DMB auxotrophy of *cobT* mutants (*25*). Both models require the existence of an alternative phosphoribosyltransferase, the expression of which requires a functional CobB protein (*26*).

In the first model, CobT is a multifunctional enzyme. It synthesizes α-ribazole-5′-phosphate from nicotinate mononucleotide (NaMN) and DMB and also participates in the biosynthesis of DMB. If this model were correct, then *cobT* mutants would not be able to synthesize DMB; however, α-ribazole-5′-phosphate could still be synthesized by CobB when exogenous DMB were provided in the medium. This would account for the DMB auxotrophy of *cobT* mutants. The weakness of this hypothesis is that biosynthesis of DMB from FMN requires at least eight enzymatic steps and it is difficult to imagine how could this be accomplished by a single enzyme. In addition, there is no in vitro evidence supporting the involvement of CobT in DMB biosynthesis.

In the second model, CobT is involved only in the catalysis of the reaction between NaMN and DMB. In addition, the alternative phosphoribosyltransferase (presumably CobB) has a higher K_M for DMB than CobT. Therefore, in *cobT* mutants, endogenous levels of DMB are not high enough for CobB function and need to be supplemented for the synthesis of α-ribazole-5′-phosphate by CobB. Therefore, this model would also account for the DMB auxotrophy of CobT mutants. At present, there is no definitive evidence in support of either hypothesis.

As mentioned earlier, the lower ligand is somewhat variable depending on the organism (*2*). This variability is further accentuated by the ability of CobT to phosphoribosylate a variety of bases including benzimidazole, 4,5-dimethyl-1,2-phenylenediamine,imidazole, histidine, adenine, and guanine (*11*). The structural basis for this lack of specificity is unknown. Likewise little is known about the mechanism of this enzyme or the amino acids that define the active site. In an effort to clarify the role of CobT in the cobalamin biosynthetic pathway and to address the questions concerning enzyme specificity and mechanism, we report here the structures of nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase as a complex with DMB and with the products of catalysis, α-ribazole-5′-phosphate and nicotinic acid, to 1.9 Å resolution in both cases.

METHODS

Protein Purification. CobT was overexpressed in strain JE2461. All purification procedures were carried out at 4 °C. Approximately 29 g of cells were resuspended in 410 mL of 50 mM Tris·HCl, pH 7.5 at 4 °C, containing 1 mM ethylenediamine-tetraacetic acid (EDTA) and 0.5 mM phenylmethanesulfonyl fluoride and disrupted by sonication. Cellular debris was removed by centrifugation at 44 000*g* for 1 h. Finely ground ultrapure ammonium sulfate (ICN Biomedicals) was added to the cell-free extract to 30% saturation and clarified by centrifugation. Finely ground ultrapure ammonium sulfate was added again to the resulting supernatant to 50% saturation and the precipitate was collected by centrifugation. The pellet was resuspended in 50 mL of 50 mM Tris·HCl, pH 7.5 and dialyzed against the same buffer at 4 °C overnight. The dialysate was clarified and loaded onto a 50 mL bed volume DEAE-Sepharose Fast Flow column (Sigma) equilibrated with 50 mM Tris·HCl, pH 7.5. The protein was eluted with a gradient from of 50 mM Tris·HCl, pH 7.5 to 0.5 M NaCl in the same buffer. CobT eluted between 0.1−0.2 M NaCl. Thereafter, dye ligand chromatography was used to further purify the protein. The pooled fractions were directly loaded onto Matrex Gel Red-A column (Amicon) equilibrated with 50 mM Tris·HCl, pH 7.5 and eluted with a gradient from 50 mM Tris·HCl, pH 7.5 to 1 M KH_2PO_4/K_2HPO_4 pH 7.5, containing 50 mM D-ribose. CobT eluted between 0.45 and 0.55 M PO_4^{3-}. The fractions were analyzed by SDS-PAGE and stained with Coomassie Blue. The fractions containing CobT were concentrated in a Centriprep-30 and dialyzed against 20 mM Tris·HCl, pH 7.5, containing 100 mM NaCl. Typically, 140 mg of protein was obtained from ∼29 g of cells.

Crystallization and X-ray Data Collection. The crystals of CobT employed for the structural investigation were grown with the hanging-drop vapor-diffusion technique. Equal volumes of protein at 6 mg/mL protein in its final storage buffer and a precipitant containing 1.3 M $NH_4H_2PO_4/(NH_4)_2HPO_4$, 1 mM DMB, and 4% poly(ethylene glycol) 400 at pH 6.0 were mixed and suspended over the precipitant solution at 4 °C. Crystals grew spontaneously and achieved sizes of 0.6 mm × 0.6 mm × 0.1 mm in 2 weeks. Precession photography revealed that the crystals belong to the space group $P2_12_12$ with unit cell dimensions of $a = 72.1$ Å, $b = 90.2$ Å, and $c = 47.5$ Å. The crystal lattice contains one subunit per asymmetric unit with a solvent content of 42%. For X-ray data collection and preparation of heavy atom derivatives, the crystals were transferred to a synthetic mother liquor of 1.0 M NaH_2PO_4/K_2HPO_4 at pH 6.0. Crystals of CobT grown in the presence of DMB (Aldrich Chemical Co.) were soaked in 1.0 M NaH_2PO_4/K_2HPO_4 at pH 6.0, containing 21 mM NaMN for 1 week. Unexpectedly, NaMN reacted with the DMB bound to the enzyme to give α-ribazole-5′-phosphate and nicotinate, thus yielding the structure of the enzyme−products complex.

X-ray data were collected to 1.9 Å resolution at 0.5 °C with a Siemens HiStar area detector at a crystal to detector distance of 12 cm. Diffraction maxima were observed to a resolution of 1.8 Å. Cu Kα radiation was generated by a Rigaku RU200 X-ray generator operated at 50 kV and 90 mA and equipped with a set of 10 cm double focusing mirrors (Charles Supper Co). Diffraction data frames, of width 0.15°, were recorded for 60 or 90 s. The frames were processed with XDS (*27*, *28*) and internally scaled with XSCALIBRE (*29*). The diffraction data statistics for the native, heavy atom derivative data, and NaMN soak are given in Table 1.

Structural Determination. The three-dimensional structure of the CobT−DMB complex was determined by multiple isomorphous replacement. Two heavy atom derivatives were prepared by soaking native crystals in solutions of 0.4 mM K_2PtCl_4 and 2.0 mM $KAu(CN)_2$. The heavy atom binding sites were located by visual inspection of difference Patterson maps and placed on a common origin with appropriate difference Fourier maps. The correct hand of the heavy atom constellation was chosen based on the sign of the anomalous signal in the heavy atom derivatives. The occupancies and positions of all heavy atom sites were refined with the program SOLVE (*30*, *31*) and the electron density map was improved with the program DM (*32*). The statistics for the

Table 1: Data Collection Statistics and Heavy Atom Refinement Statistics

	native[a] CobT−DMB	$KAu(CN)_2$[b]	K_2PtCl_4[b]	CobT−α-ribazole-5′-phosphate nicotinate[a]
concentration (mM)		2.0	0.5	
length of soak (h)		21	13	
number of crystals	2	1	1	2
resolution (Å)	1.88	2.2	2.2	1.9
average I/σ	13.6 (2.9)	10.7 (2.7)	10.6 (2.4)	10.9 (3.0)
unique reflections	24 807	14 498	15 218	22 283
redundancy	3.9 (1.7)	2.1 (1.4)	2.1 (1.3)	3.1 (1.3)
completeness (%)	90.8 (64.7)	82.1 (46.6)	77.1 (48.4)	87.3 (65.0)
R_{merge}[c]	6.5	6.4	5.6	6.6
R_{iso}[d]		15.6	13.9	
number of sites		2	1	
R_{Cullis}[e]		53	63	
phasing power[f]		1.25 (1.41)	0.93 (1.07)	
figure of merit	0.51			

[a,b] The values in parentheses are from resolution shells of 1.98−1.88 and 2.31−2.20, respectively. [c] $R_{merge} = \sum[|I_{hi}| - |I_h|]/\sum I_{hi} \times 100$ where I_{hi} and I_h are the intensities of individual and mean structure factors. [d] $R_{iso} = \sum[|F_h| - |F_n|]/\sum_n F_n \times 100$ where F_h and F_n are the heavy atom and native structure factors. [e] $R_{Cullis} = \sum[|F_h - F_n| - |F_{hc}|]/\sum|F_h - F_n| \times 100$ where the summation is carried out over the centric data only. F_h and F_n are the observed heavy atom derivative and native structure factors and F_{hc} is the calculated structure for the heavy atoms alone. [f] The phasing power is defined as the mean value of the heavy atom structure factor divided by the lack-of-closure error.

Table 2: Least Squares Refinement Statistics[a]

	CobT−DMB	CobT·α-ribazole-5′-phosphate nicotinate
resolution limits (Å)	30−1.9	30−1.9
final R_{factor}[b]	17.2	16.3
R_{free}[c]		24.9
number of reflections used	19 486	21 884
number of protein atoms	2418	2418
number of solvent molecules	132	158
other molecules, ions	1 DMB	1 α-ribazole-5′-phosphate
	1 phosphate ion	1 nictotinic acid
average *B* values Å^2		
main chain atoms	16.5	15.5
all protein atoms	19.9	18.8
solvent atoms	29.1	30.8
weighted *rms* deviations from ideality		
bond lengths (Å)	0.009	0.010
bond angles (deg)	1.86	1.83
planarity (trigonal) (Å)	0.002	0.003
planarity (others) (Å)	0.010	0.011
torsional angles (deg)[d]	15.6	15.7

[a] TNT refinement. [b] $R = \sum||F_o| - k|F_c||/\sum|F_o|$. [c] R_{factor} for 10% of the data excluded from the refinement. The final working R_{factor} was 16.0%. During the final cycles of refinement all data was included in the refinement whereupon the working R_{factor} rose to 16.3%. The statistics quoted refer to all of the data. [d] No restraints were placed on torsional angles during refinement.

heavy atom refinement are given in Table 1. The initial model was built with the program FRODO (*33*) on an Evans and Sutherland PS390 and subsequently with TURBO (*34*). The model was refined by the method of least squares with the program TNT (*35, 36*). During the early stages of the model building, the heavy atom phases were combined with model phases, utilizing SIGMAA weighting (*37*). During the later stages of refinement, solvent molecules were added in locations where the electron density and geometry were consistent with a water molecule with the program PEKPIK in the CCP4 package (*38*). DMB and a phosphate ion were built into unequivocal electron density. The final *R*-factor for the CobT−DMB complex is 17.2% for all measured X-ray data from 30 to 1.9 Å with root-mean-square deviations from ideal geometry of 0.009 Å for bond lengths, 1.9° for bond angles, and 0.002 Å for groups of atoms expected to be coplanar. R_{free} was not used in the structure determination of the CobT−DMB complex. Analysis of the backbone dihedral angles with PROCHECK revealed that for the CobT−DMB complex 92% of the residues conformed to the most favorable regions where the other 3% residues conformed to other additionally allowed regions and three residues, Asp^{263}, Arg^{314}, and Glu^{317}, lie in the "generously allowed regions" (*39*). These three residues, which all have well-defined electron density, are associated with the substrate binding pocket. Refinement statistics are given in Table 2. A representative portion of the electron density map is shown in Figure 3a.

CobT crystals grown in the presence of DMB were soaked in 21 mM NaMN and the data were collected as before (Table 1). Refinement of the products complex started from the final model of CobT−DMB from which all water molecules, DMB, and phosphate ion were omitted. This starting model was refined against a working data set that contained a random 90% of the data. The resultant $2F_o - F_c$ and $F_o - F_c$ electron density maps revealed the location of α-ribazole-5′-phosphate and nicotinic acid. These molecules were included in the model and refined with the

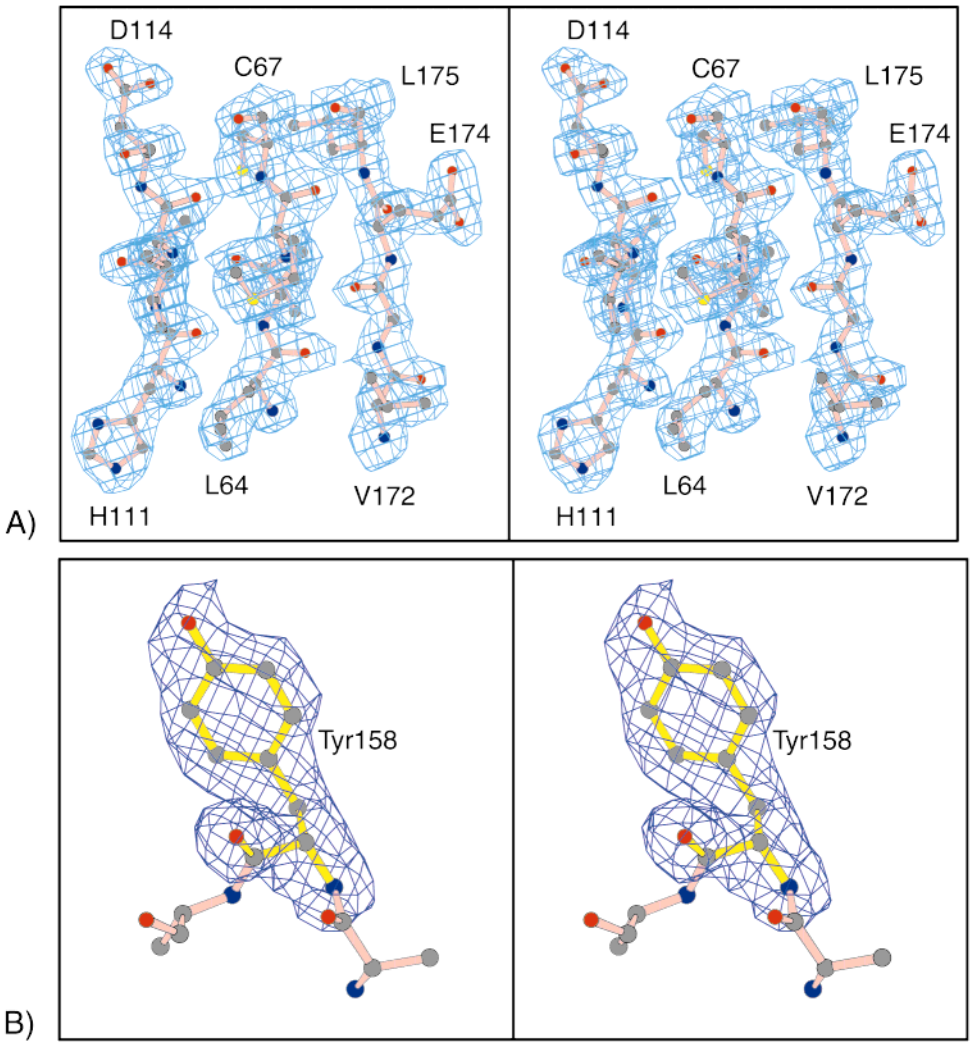


FIGURE 3: Stereoview of representative electron density in the CobT−DMB complex. (A) Electron density associated with the large β-sheet calculated with coefficients of the form $2F_o - F_c$, and (B) Electron density for Tyr^{158} where this residue was omitted from the refinement and phase calculation. The map was calculated with coefficients of the form $F_o - F_c$. Figures 3, 6, and 7 were prepared with the programs MOLSCRIPT and MOLDED (*54*, *55*).

program TNT to give an initial R_{factor} of 20.4%. Further manual adjustment of the model and least squares refinement together with the addition of water molecules lowered the R_{factor} to 16.0% with an R_{free} of 24.9%. The final model was obtained by refinement against the complete data set to yield a R_{factor} of 16.3%. The rms difference in the positions of the α-carbons in CobT−DMB and the CobT−products complex is 0.12 Å. Analysis of the backbone dihedral angles with PROCHECK revealed that for the CobT−products complex 92% of the residues conformed to the most favorable regions where the other 7% residues conformed to additionally allowed regions (*39*). Three residues lie in the "generously allowed regions" as observed in the CobT−DMB complex. The distribution of dihedral angles is essentially identical to that found in the CobT−DMB complex. The refinement statistics are given in Table 2. The X-ray coordinates have been deposited in the Protein Data Bank with file names 1D0S and 1D0V for the complex of CobT with 5,6-dimethybenzimidazole and the reaction products, respectively.

RESULTS AND DISCUSSION

The final model of CobT complexed with DMB contains 346 amino acid residues out of 356 expected from the sequence together with one DMB molecule, one phosphate ion, and 132 water molecules. The polypeptide chain extends without breaks from Leu^4 to Pro^{349}. The average temperature factor for all protein atoms is 19.9 Å^2, for protein main chain atoms is 16.5 Å^2, while for solvent molecule is 29.1 Å^2 (Table 2). Several protein residues of the final model are poorly ordered in the structure. The side chain of the residues His^5, Gln^{24}, Lys^{109}, Leu^{146}, Arg^{157}, Asp^{161}, Arg^{165}, Lys^{196}, Glu^{197}, Ile^{201}, Pro^{207}, Arg^{209}, Asp^{215}, Arg^{229}, Arg^{297}, and Glu^{306} are disordered and have been built as alanine residues in the current model. Most of these residues are located at the protein surface. The electron density of residue 158 suggested that this residue is a tyrosine rather than a cysteine as was indicated in the original gene sequence (Figure 3b). Since the genetic code specifying a cysteine and tyrosine differ only in one base (UGU or UGC for a cysteine and UAU or UAC for a tyrosine) the sequence was redetermined. This confirmed that this residue is indeed a tyrosine. CobT contains eight cysteine residues. Two of these, Cys^{160} and Cys^{256}, form a disulfide bond that links two adjacent α-helices.

Tertiary structure of CobT. The structure of one protomer of CobT is composed of a six-stranded parallel β-sheet and a total of 17 α-helices (Figures 4 and 5). It can be divided

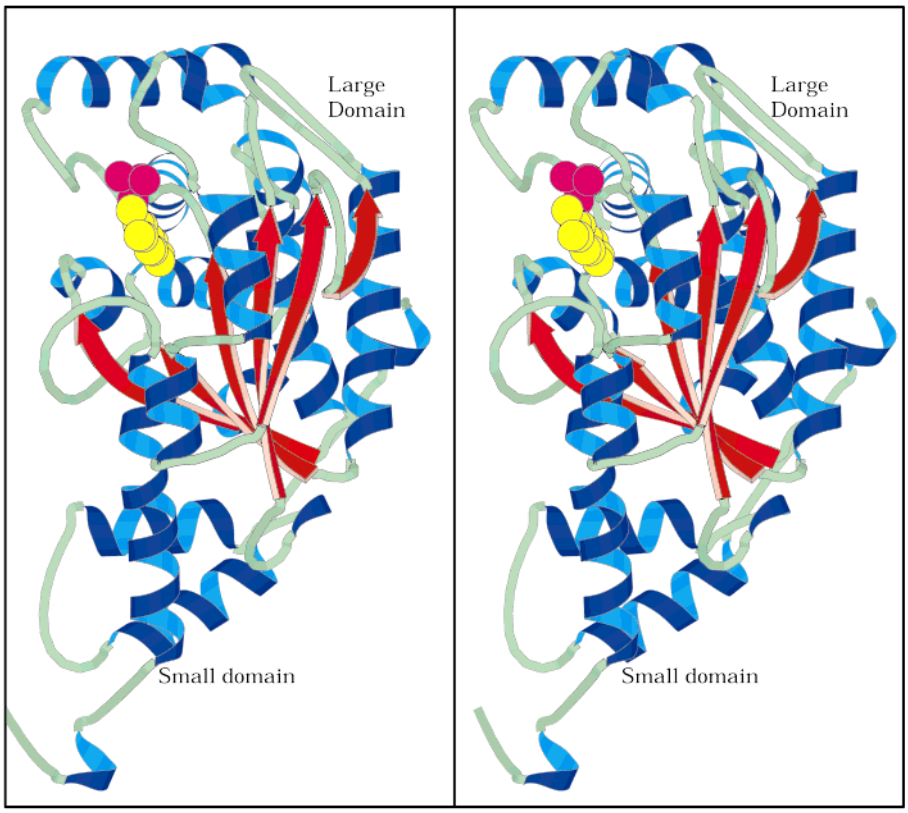


FIGURE 4: Stereo ribbon representation of the CobT monomer with DMB and phosphate bound. Figures 4 and 5 were prepared with the program MOLSCRIPT (*54*).

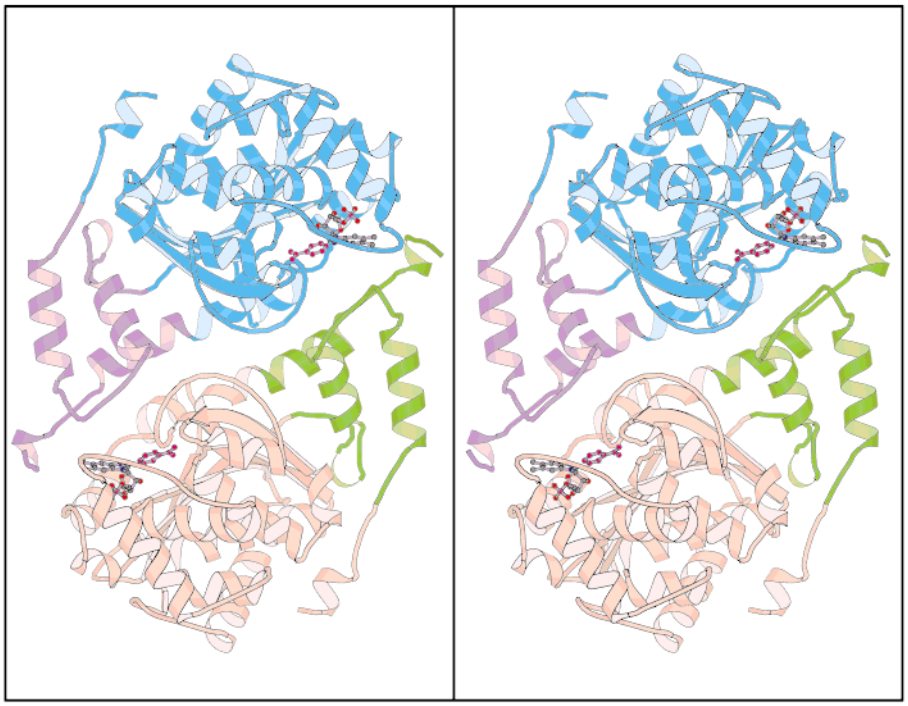
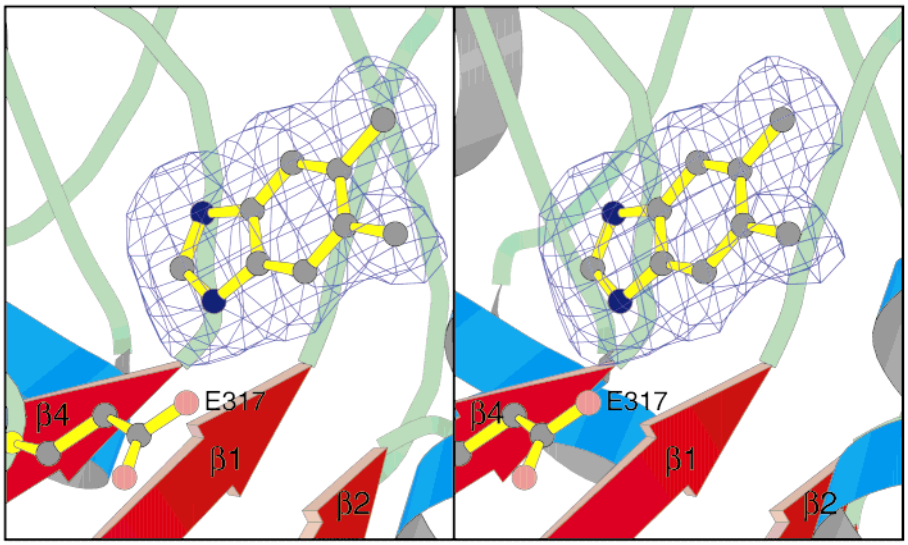
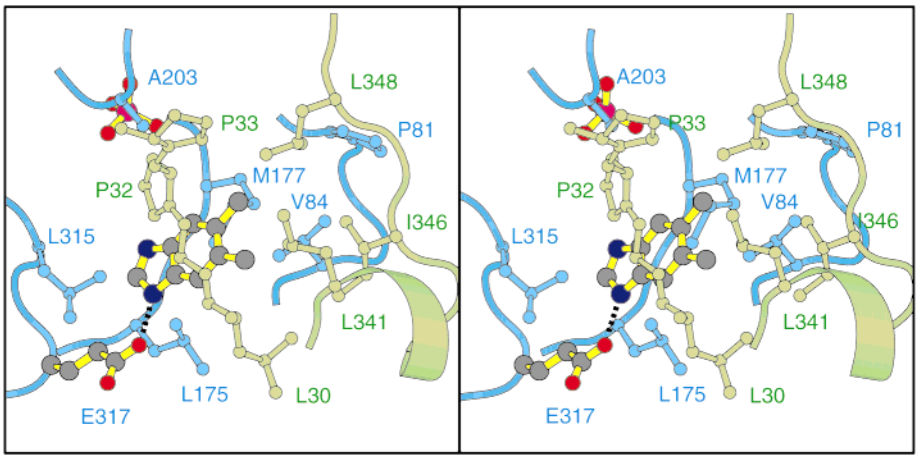


FIGURE 5: Stereo ribbon representation of the CobT dimer viewed along the crystallographic 2-fold axis. The large domains are depicted in blue and pink whereas the small domains are colored in magenta and green.

approximately into two domains. The large domain extends from Gly^{59} through Ala^{331} and is built from six strands of parallel β-sheet and surrounding α-helices. The small domain contains residues from both the N-terminal and C-terminal sections of the polypeptide chain. The first eleven residues at the N-terminus of the protein lie against the large domain, then the polypeptide chain proceeds into two antiparallel helices and a connecting loop that form the bulk of the small domain. Part of a helix that joins the large and small domain together with a section of random coil are contributed by the C-terminal section of the protein that completes the small domain. The major tertiary feature of the small domain is a



(a)



(b)

FIGURE 6: Stereoview of the difference electron density for (a) DMB and (b) a close up view of the hydrophobic binding pocket for this substrate. The electron density was calculated from coefficients of the form $F_o - F_c$ where the substrate was omitted from the refinement and phase calculation. This figures shows that the binding pocket is built with residues from the large domain of one subunit (blue) and the small domain of the 2-fold related subunit (green).

three helical bundle. Within the large domain, the order of the strands in the sheet is β6, β5, β4, β1, β2, and β3, which is the arrangement observed in the Rossmann fold found in many dinucleotide binding proteins. At first sight, this topological feature is consistent with the function of CobT to bind nicotinate mononucleotide; however, as discussed later, the orientation and position of the α-ribazole-5′-phosphate is different from that expected based on a dinucleotide binding motif. The exact structural motif found in CobT does not appear to be related to any other fold in the protein data bank as determined with the program DALI (*40*).

As shown in Figure 5, the quaternary structure of CobT is a dimer where the two subunits are related by crystallographic 2-fold symmetry. The overall size of the dimer is approximately 40 × 55 × 80 Å where the molecule is elliptical when viewed along the crystallographic 2-fold axis, but bilobal when viewed perpendicular to this axis. The total buried surface area of the dimer is about 2300 Å^2/subunit as determined with the program GRASP (*41*) and represents about 19% of the surface area of the monomer. This is typical for dimeric proteins of this molecular weight (*42*) and is indicative of an extensive interface between the two subunits. A substantial section of the interface is provided by the small domain and the final helix in the large subunit. This interface is important for the biological function of this enzyme, since the DMB binding site is formed from components of both subunits.

The electron density for DMB in the structure of the CobT−DMB complex is unambiguous (Figure 6a). The substrate is found in a prominent hole in the molecule toward

the periphery of the dimer located adjacent to the dimer interface (Figures 5 and 6b) and involves the C-terminal domain of the second subunit. The imidazole component of the polyaromatic ring points in toward the C-terminal end of β-strand 4, whereas the two methyl groups are directed toward the subunit interface. The binding site of DMB is part of a prominent cavity that is occupied by a phosphate ion and several water molecules. This cavity becomes the binding site for α-ribazole-5′-phosphate and nicotinic acid in the products complex where the site of the phosphate ion becomes occupied by the phosphate moiety of α-ribazole-5′-phosphate. The active site cavity is nestled between the loops at the C-terminal ends of β-strands 4 and 5 and is enclosed by extension of loops at the ends of β-strands 1 and 6 and is flanked on one side by a helix between Ile^{83} and Met^{92}. In addition, the C-terminal domain of the neighboring subunit contributes two loops from $Leu^{30'}$ to $Ser^{35'}$ and $Ser^{344'}$ to $Pro^{349'}$ that play an important role in defining the binding site for DMB.

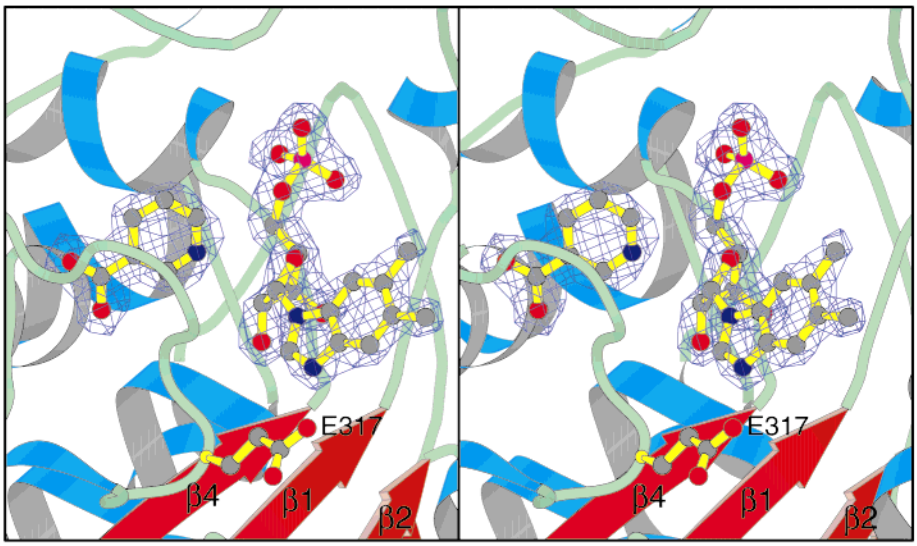
The DMB binding pocket is highly hydrophobic, which is consistent with the nature of this substrate. As shown in Figure 6b, the side chains of Pro^{81}, Val^{84}, Gln^{88}, Leu^{175}, Met^{177}, Ala^{203}, and Leu^{315} are involved in establishing this hydrophobic environment. In addition, Leu^{30}, Pro^{32}, Pro^{33}, Leu^{341}, Ile^{346}, and Leu^{348} from the neighboring subunit of the dimer make a major contribution to the hydrophobicity of the binding pocket. There are no contacts less than 3.5 Å between the dimethylbenzene moiety and any of the atoms that form the binding pocket; indeed, most of the atoms are located more than 3.8 Å from the substrate. The size and nature of the binding pocket is consistent with the ability of CobT to phosphoribosylate alternative bases such as imidazole, benzimidazole, or 4,5-dimethyl-1,2-phenylenediamine (*11*), although it is less easy to account for the utilization of adenine and histidine. There is a prominent interaction between the imidazole part of the substrate and the binding pocket. A hydrogen bond is observed between an imidazole nitrogen and Glu^{317}, which as noted earlier adopts an unfavorable set of conformational angles. Although this residue is clearly important in coordinating the substrate it may also fulfill a role as a catalytic base as discussed later. In addition to DMB, the substrate binding pocket contains a well-defined phosphate ion that is coordinated by residues in the loop between the second half of the loop that connects β-strand 4 to its succeeding α-helix between Glu^{174} and Gly^{176}. This binding site is occupied by the phosphate of the products complex and is described below.

Products Complex. The overall structure of the products complex is essentially identical to that of the CobT−DMB complex except for small differences in the number of water molecules and small changes in the disordered side chains. As might be expected, the DMB and phosphate moiety of α-ribazole-5′-phosphate are bound to the same site as DMB and phosphate of the CobT−DMB complex. The ribose ring interacts with the initial half of the loop that connects β-strand 4 to its succeeding α-helix between Glu^{174} and Gly^{176}, whereas the phosphate moiety interacts with the remainder of the loop and the first turn of the α-helix (Figure 7a). Additional coordination of the phosphate is provided by the amide hydrogen of Ala^{203} and a well-defined water network. These water molecules all have low *B*-values and are key components in the coordination of the phosphate ion. Interestingly, the side chain of Lys^{213} is pointed directly toward the phosphate moiety and is coordinated via a water molecule to one of the terminal oxygens of the phosphate. It would appear likely that Lys^{213} contributes indirectly to the neutralization of the charge on the phosphate moiety of α-ribazole-5′-phosphate.

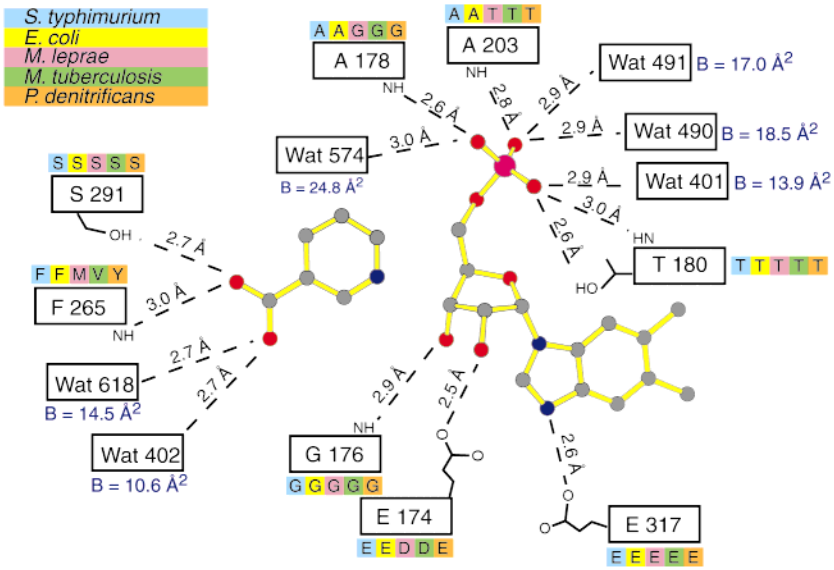
The nicotinate resides in a substantial cavity formed by the loops at the C-terminal ends of β-strands 5 and 6 and by the extensions of the helices that follow β-strands 4 and 6. The nicotinate is located such that the nitrogen of the pyridine ring is located 3.3 Å from C1 of the ribose ring and is oriented opposite to the covalent bond to DMB that was formed by its displacement (Figure 7a). There are comparatively few interactions between the nicotinate and the binding pocket. The pyridine ring is located in a hydrophobic environment provided by the main chain atoms of Arg^{314} and Leu^{315} and side chains of Leu^{266} and the Cγ of Thr^{180}. One oxygen of the carboxylate group interacts with the amide hydrogen of Phe^{265} and the hydroxyl of Ser^{291} with distances of 3.0 and 2.7 Å, respectively. Conversely, the second oxygen interacts primarily with the well-defined solvent cavity associated with the nicotinic acid binding pocket. The coordination site for the oxygen associated with Phe^{265} and Ser^{291} would not be consistent with an amino group; however, it is not clear why an amino group could not be accommodated at the site of the second oxygen. Consequently, there is no obvious structural reason CobT prefers nicotinate mononucleotide to nicotinamide mononucleotide as is observed (*11*). Interestingly, all of the residues that are involved in binding the nicotinate and the α-ribazole-5′-phosphate are well-conserved among CobT homologues from other bacteria (Figure 7b).

The orientation and location of NaMN and α-ribazole-5′-phosphate are opposite from what might be expected from the presence of a dinucleotide binding motif or Rossmann fold in CobT. In most enzymes that contain a Rossmann fold, the phosphates of the dinucleotide are located at the midpoint of the C-terminal ends of the parallel β-sheet and the nucleotide bases are positioned toward the outer edges of the sheet. By contrast, in CobT α-ribazole-5′-phosphate is oriented such that the phosphate moiety is located at the periphery of the sheet whereas the base is found adjacent to the midpoint. This orientation places the active site close to the center of the β-sheet, whereas in dehydrogenases the active site typically lies at the edge of a sheet and is associated with a second domain. Although this is unexpected, it reflects the versatility of such commonly observed structural motifs.

Mechanistic Considerations. It is known from the stereochemistry of the reactants and products that the reaction catalyzed by CobT proceeds by direct displacement since the NaMN and α-ribazole-5′-phosphate have opposite configurations at C1 of the ribose ring. The orientation of the products are consistent with this geometry as noted above. Under physiological conditions, 5,6-dimethylbenzimidazole would be neutral in solution since the pK_a's for the two nitrogens in the imidazole heterocycle are approximately 5 and 13. Conversely, the nicotinate product will most likely exist in its deprotonated state, since the pK_a for the heterocyclic nitrogen is approximately 5.0 (*43*). Thus, during the phosphoribosyl transfer reaction a proton must be abstracted from the DMB in order to generate a neutral



(a)



(b)

FIGURE 7: Stereoview of (a) the difference electron density for α-ribazole-5′-phosphatephosphate and nicotinic acid in CobT−products complexes and (b) schematic representation of the interactions between CobT and the products complex. The electron density was calculated from coefficients of the form $F_o - F_c$ where both of these molecules were excluded from the phase calculations and refinement. In (b) the residues in CobT from other species that are equivalent to those that interact with the products complex are shown. The sequences and accession numbers used for this alignment were; *S. typhimurium*, Q05603; *P. denitrificans*, P29935; *M. leprae*, CAB11373; *E. coli*, P36562; *M. tuberculosis* Q05603. The alignment was performed with the GCG program package (*56*).

α-ribazole moiety. It seems likely that Glu^{317} fulfils the role of a catalytic base, since it forms a short hydrogen bond (2.6 Å) between Oϵ2 of Glu^{317} and N3 of α-ribazole-5′-phosphate. Interestingly, in the products complex, the distance between these atoms is still only 2.6 Å, which suggests that the proton remains associated with the hydrogen bond and most likely still resides on N3 of α-ribazole-5′-phosphate.

CobT prefers nicotinate mononucleotide to nicotinamide mononucleotide by a factor of ~44 in K_M where there is little change in k_{cat} (*11*). This preference is contrary to the pK_a's of the heterocyclic nitrogens of nicotinate and nicotinamide, which are 4.8 and 3.4, respectively. Thus, it would be expected that nicotinamide would be a better leaving group than nicotinate. As noted before, the binding pocket for nicotinate does not show any obvious reason it should prefer this substrate; however, it is conceivable that the specificity is controlled by the orientation of the water molecules in the active site that are well-defined in this instance.

CONCLUSIONS

The reaction catalyzed by CobT is intrinsically similar to that performed by a larger class of phosphoribosyltrans-

ferases. In most instances, the phosphophoribosyl group is transferred from α-phosphoribosylpyrophosphate (PRPP) to a nitrogenous base with an exchange of configuration at C1 of the ribose and loss of pyrophosphate. With CobT, the source of the phosphoribosyl moiety is NaMN that is itself in the β-configuration. Consequently, the product of the reaction catalyzed by CobT is α-ribazole-5′-phosphate, which is the configuration found in the nucleotide loop in cobalamin. Interestingly, NaMN is itself synthesized by a PRPP dependent phosphoribosyltransferase. Consequently, it might be expected that CobT would share some similarities with a broader class of phosphoribosyltransferases where the dimethylbenzimidazole would be chemically equivalent to the pyrophosphate moiety in the transfer reaction.

In recent years, structures for a significant group of phosphoribosyltransferases (PRTase) have been determined. This list includes hypoxanthine-guanine phosphoribosyltransferase (*44*), orotate PRTase(*12, 45, 46*), xanthine PRTase (*47*), quinolinic PRTase (*13*), glutamine PRTase amidotransferase (*48*), and uracil PRTase-transcription attenuator from *Bacillus subtilis* (*49*) where the source of the phosphoribosyl moiety is always phosphoribosylpyrophosphate. These proteins fall into two distinct structural groups as represented by orotate phosphoribosyltransferase and quinolinic phosphoribosyltransferase (*13*), which contain two very different protein folds. The type I fold observed in orotate phosphoribosyltransferase is the most abundant motif and is dominated by a central parallel five-stranded β-sheet surrounded by α-helices. This fold includes a common motif of thirteen residues that is involved in recognizing phosphoribosylpyrophosphate. Interestingly, the structure of CobT is also based on a parallel β-sheet; however, the connectivity of the β-strands is quite different from that of the type I phosphoribosyltransferases. The type II fold observed only in quinolinic phosphoribosyltransferase is completely different and consists of two domains; an N-terminal open-faced sandwich domain and a C-terminal seven-stranded α/β barrel. The active site is associated with the C-terminal end of the α/β barrel. The structure of CobT described here establishes that there is another distinct fold capable of catalyzing phosphoribosyl transfer.

Although the topology of the type I phosphoribosyltransferases and CobT are different, in both cases the substrates are bound by the loops at the C-terminal ends of the β-strands. Even so, the mode of binding and orientation of the substrates are quite different. An extensive series of studies on the type I phosphoribosyltransferases have demonstrated that these enzymes undergo a significant conformational change during the catalytic cycle. Typically, loops that are solvent exposed in the absence of substrates fold down and create an active site only when all of the substrates are bound (*50*−*52*). In the case of CobT, it would appear that no major changes occur when NaMN binds in the active site. Comparison of CobT with the type I phosphoribosyltransferases shows that the orientation of the nitrogenous product in the type I enzymes is unrelated to the position of NaMN and α-ribazole-5′-phosphate in CobT. This suggests that CobT represents a unique form of phosphoribosyl transferase and does not share an evolutionary ancestry with the PRPP dependent enzymes.

Finally, given the structure of CobT, it is difficult to envision how this enzyme can be solely responsible for the conversion of FMN to DMB as proposed by Chen et al. (*18*). Even if CobT were able to catalyze this conversion, it is unlikely that it would occur via the pathway proposed in the literature (*21*), mainly because there are eight proposed reactions in this pathway, several of which involve substantially different chemistry. It is formally possible, however, that CobT may be involved in some step in the conversion of FMN to DMB. This possibility is under investigation.

ACKNOWLEDGMENT

We thank Dr. Jodi Trzebiatowski for the initial protein samples and helpful discussions throughout this investigation. We also thank Drs. Cary Bauer and Andy Gulick for help with heavy atom refinement and model building.

REFERENCES

1. Folkers, K. (1982) in B_{12} (Dolphin, D., Ed.) pp 1, Wiley, New York.
2. Rondon, M. R., Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1997) *Prog. Nucleic Acid Res. Mol. Biol. 56*, 347−384.
3. Lenhert, P. G., and Hodgkin, D. C. (1961) *Nature 192*, 937−938.
4. Scott, A. I. (1993) *Angew. Chem., Int. Ed.* (*Engl.*) *32*, 1223−1243.
5. Blanche, F., Cameron, B., Crouzet, J., Debussche, L., Thibaut, D., Vuilhorgne, M., Leeper, F. J., and Battersby, A. R. (1995) *Angew. Chem., Int. Ed.* (*Engl.*) *34*, 383−411.
6. Roth, J. R., Lawrence, J. G., and Bobik, T. A. (1996) *Annu. Rev. Microbiol. 50*, 137−81.
7. O'Toole, G. A. (1994), University of Wisconsin, Madison, Ph.D. thesis.
8. O'Toole, G. A., Rondon, M. R., and Escalante-Semerena, J. C. (1993) *J. Bacteriol. 175*, 3317−3326.
9. Roth, J. R., Lawrence, J. G., Rubenfield, M., Kieffer-Higgins, S., and Church, G. M. (1993) *J. Bacteriol. 175*, 3303−3316.
10. O'Toole, G. A., Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1994) *J. Biol. Chem. 269*, 26503−26511.
11. Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1997) *J. Biol. Chem. 272*, 17662−17667.
12. Scapin, G., Grubmeyer, C., and Sacchettini, J. C. (1994) *Biochemistry 33*, 1287−94.
13. Eads, J. C., Ozturk, D., Wexler, T. B., Grubmeyer, C., and Sacchettini, J. C. (1997) *Structure 5*, 47−58.
14. Bazan, J. F., Fletterick, R. J., and Pilkis, S. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9642−9646.
15. Tauler, A., Lin, K., and Pilkis, S. J. (1990) *J. Biol. Chem. 265*, 15617−15622.
16. Pilkis, S. J., Claus, T. H., Kurland, I. J., and Lange, A. J. (1995) *Annu. Rev. Biochem. 64*, 799−835.
17. Maggio-Hall, L. A., and Escalante-Semerena, J. C. (1999) *Proc. Natl. Acad. Sci. U.S.A., 96*, 11798−11803.
18. Chen, P., Ailion, M., Weyand, N., and Roth, J. (1995) *J. Bacteriol. 177*, 1461−9.
19. Renz, P. (1970) *FEBS Lett. 6*, 187−189.
20. Renz, P., and Weyhenmeyer, R. (1972) *FEBS Lett. 22*, 124−126.
21. Horig, J. A., and Renz, P. (1980) *Eur. J. Biochem. 105*, 587−92.
22. Keck, B., Munder, M., and Renz, P. (1998) *Arch. Microbiol. 171*, 66−8.
23. Escalante-Semerena, J. C., Johnson, M. G., and Roth, J. R. (1992) *J. Bacteriol. 174*, 24−9.
24. Trzebiatowski, J. R. (1998) in *Bacteriology*, University of Wisconsin, Madison, Ph.D. thesis.
25. Trzebiatowski, J. R., O'Toole, G. A., and Escalante-Semerena, J. C. (1994) *J. Bacteriol. 176*, 3568−3575.
26. Tsang, A. W., and Escalante-Semerena, J. C. (1998) *J. Biol. Chem. 273*, 31788−94.
27. Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 67−71.
28. Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 916−924.
29. Wesenberg, G., and Rayment, I. (1997) In preparation.

30. Terwilliger, T. C., and Eisenberg, D. (1983) *Acta Crystallogr. A39*, 813−817.
31. Terwilliger, T. C. (1997) in *Methods in Enzymology* (Carter, C. W. J., Sweet, R. M., Abelson, J. N., and Simon, M. I., Eds.), *276*, pp 530−537, Academic Press, New York.
32. Cowtan, K., and Main, P. (1998) *Acta Crystallogr. D54*, 487−493.
33. Jones, T. A. (1985) in *Methods in Enzymology* (Wycoff, H. W., Hirs, C. H. W., and Timasheff, S. N., Eds.), *115*, pp 157−171, Academic Press Inc., New York.
34. Roussel, A., and Cambillau, C. (1991) in *Silicon Graphics Geometry Partners Directory86*, Silicon Graphics.
35. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489−501.
36. Tronrud, D. E. (1997) *Methods Enzymol. 277*, 306−19.
37. Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
38. CCP4 (1994) *Acta Crystallogr. D50*, 760−763.
39. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
40. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−38.
41. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins: Struct. Funct., Genet. 11*, 281−296.
42. Miller, S., Lesk, A. M., Janin, J., and Chothia, C. (1987) *Nature 328*, 834−836.
43. Fasman, G. (1976) *Physical and Chemical Data*, *Vol. 1*.
44. Eads, J. C., Scapin, G., Xu, Y., Grubmeyer, C., and Sacchettini, J. C. (1994) *Cell 78*, 325−34.
45. Ozturk, D. H., Dorfman, R. H., Scapin, G., Sacchettini, J. C., and Grubmeyer, C. (1995) *Biochemistry 34*, 10764−70.
46. Scapin, G., Ozturk, D. H., Grubmeyer, C., and Sacchettini, J. C. (1995) *Biochemistry 34*, 10744−54.
47. Vos, S., de Jersey, J., and Martin, J. L. (1997) *Biochemistry 36*, 4125−34.
48. Krahn, J. M., Kim, J. H., Burns, M. R., Parry, R. J., Zalkin, H., and Smith, J. L. (1997) *Biochemistry 36*, 11061−8.
49. Tomchick, D. R., Turner, R. J., Switzer, R. L., and Smith, J. L. (1998) *Structure 6*, 337−50.
50. Focia, P. J., Craig, S. P., 3rd, and Eakin, A. E. (1998) *Biochemistry 37*, 17120−7.
51. Balendiran, G. K., Molina, J. A., Xu, Y., Torres-Martinez, J., Stevens, R., Focia, P. J., Eakin, A. E., Sacchettini, J. C., and Craig, S. P., 3rd (1999) *Protein Sci. 8*, 1023−31.
52. Shi, W., Li, C. M., Tyler, P. C., Furneaux, R. H., Grubmeyer, C., Schramm, V. L., and Almo, S. C. (1999) *Nat. Struct. Biol. 6*, 588−93.
53. Thompson, T. B., Thomas, M. G., Escalante-Semerena, J. C., and Rayment, I. (1998) *Biochemistry 37*, 7686−7695.
54. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
55. Fisher, A. J. (1996), copyright University of Wisconsin, Madison, WI.
56. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acid Res. 12*, 387−395.

BI991752C