THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 40, Issue of October 5, pp. 37612–37620, 2001

Printed in U.S.A.

Structural Investigation of the Biosynthesis of Alternative Lower Ligands for Cobamides by Nicotinate Mononucleotide: 5,6-Dimethylbenzimidazole Phosphoribosyltransferase from *Salmonella enterica**

Received for publication, June 11, 2001, and in revised form, July 2, 2001
Published, JBC Papers in Press, July 5, 2001, DOI 10.1074/jbc.M105390200

Cheom-Gil Cheong‡, Jorge C. Escalante-Semerena§¶, and Ivan Rayment‡‖

From the Departments of ‡Biochemistry and §Bacteriology, University of Wisconsin, Madison, Wisconsin 53706

Nicotinate mononucleotide (NaMN):5,6-dimethylbenzimidazole phosphoribosyltransferase (CobT) from *Salmonella enterica* plays a central role in the synthesis of α-ribazole, a key component of the lower ligand of cobalamin. Surprisingly, CobT can phosphoribosylate a wide range of aromatic substrates, giving rise to a wide variety of lower ligands in cobamides. To understand the molecular basis for this lack of substrate specificity, the x-ray structures of CobT complexed with adenine, 5-methylbenzimidazole, 5-methoxybenzimidazole, *p*-cresol, and phenol were determined. Furthermore, adenine, 5-methylbenzimidazole, 5-methoxybenzimidazole, and 2-hydroxypurine were observed to react with NaMN within the crystal lattice and undergo the phosphoribosyl transfer reaction to form product. Significantly, the stereochemistries of all products are identical to those found *in vivo*. Interestingly, *p*-cresol and phenol, which are the lower ligand in *Sporomusa ovata*, bound to CobT but did not react with NaMN. This study provides a structural explanation for how CobT can phosphoribosylate most of the commonly observed lower ligands found in cobamides with the exception of the phenolic lower ligands observed in *S. ovata*. This is accomplished with minor conformational changes in the side chains that constitute the 5,6-dimethylbenzimidazole binding site. These investigations are consistent with the implication that the nature of the lower ligand is controlled by metabolic factors rather by the specificity of the phosphoribosyltransferase.

Cobalamin is the largest and most complex cofactor found in biological systems (Fig. 1). It consists of a corrin ring, which provides four of the ligands to the central cobalt ion, together with an upper and lower ligand that complete the coordination sphere. In its biologically functional form, the upper ligand of the coenzyme is an adenosyl group that is covalently attached to the cobalt. This labile bond cobalt-carbon is the source of the unusual chemistry facilitated by cobamides. In contrast to the upper ligand, the lower ligand shows considerable variation (3).

The range of lower ligands that have been observed is extensive and depends on the nature of the microorganism (Table I). The list includes 5,6-dimethylbenzimidazole (DMB),[1] 5-methylbenzimidazole, 5-methoxybenzimidazole, 5-hydroxybenzimidazole, 5-methoxy-6-methylbenzimidazole, adenine, 2-methylsulfinyladenine, 2-methylsulfonyladenine, *p*-cresol, and phenol (3, 4). Of these, 5,6-dimethylbenzimidazole is the most commonly observed base of the lower ligand and is utilized by organisms such as *Salmonella enterica* and *Pseudomonas denitrificans*. These ligands fall into two groups depending on whether a carbon-nitrogen or carbon-oxygen bond is formed to the ribose.

The enormous variation in the lower ligand, even within a single organism, presents a biochemical problem of how the variability of this essential component of the cofactor is accommodated by cobamide-dependent enzymes. Structural studies on both methionine synthase and methylmalonyl-CoA mutase have shown that the lower ligand of cobalamin is replaced by a histidine side chain (5, 6). As a consequence, a wide variety of lower ligands could be accommodated by these enzymes. Interestingly, in diol dehydratase, the lower ligand remains coordinated to the corrinoid (7), which implies that this enzyme should be more selective for its cofactor, as is observed (8, 9).

A related question to how enzymes can utilize a range of lower ligands is how the enzymes responsible for cobamide biosynthesis are able to accommodate such a wide range of ligands. Much less is known about how the intermediates in cobamide biosynthesis are recognized, although considerable effort has been made to understand the biosynthesis of cobalamin (10, 11). Twenty-four genes have been identified to be directly involved in the construction of this cofactor (12). Of these, *cobC*, *cobS*, *cobT*, and *cobU* in *S. enterica* are required for synthesis of the lower ligand and assembly of the nucleotide loop (Fig. 2). CobU (adenosylcobinamide kinase/adenosylcobinamide phosphate guanylyltransferase) is the central enzyme in the biosynthesis of the nucleotide loop and functions to activate adenosylcobinamide by the attachment of guanosine diphosphate to form adenosylcobinamide-GDP (13–15). CobS (cobalamin synthase) is the final enzyme in cobalamin biosynthesis and combines the activated cobinamide or adenosylcobin-

* This work was supported in part by National Institutes of Health Grants GM58281 (to I. R.) and GM40313 (to J. E.-S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1JH8 (adenine), 1JHA (α-AMP and nicotinate), 1JHM (5-methylbenzimidazole), 1JHO (reaction products of 5-methylbenzimidazole and NaMN), 1JHP (5-methoxybenzimidazole), 1JHQ (reaction products of 5-methoxybenzimidazole and NaMN), 1JHR (reaction products of 2-hydroxypurine and NaMN), 1JHU (p-cresol), 1JHV (p-cresol and NaMN), 1JHX (phenol), and 1JHY (phenol and NaMN)) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

¶ To whom correspondence may be addressed: Dept. of Bacteriology, Fred Hall, Linden Dr., Madison, WI 53706. E-mail: jcescala@facstaff.wisc.edu.

‖ To whom correspondence may be addressed: Dept. of Biochemistry, 433 Babcock Dr., Madison, WI 53706. Tel.: 608-262-0437; Fax: 608-262-1319; E-mail: Ivan_Rayment@biochem.wisc.edu.

[1] The abbreviations used are: DMB, 5,6-dimethylbenzimidazole; NaMN, nicotinate mononucleotide.

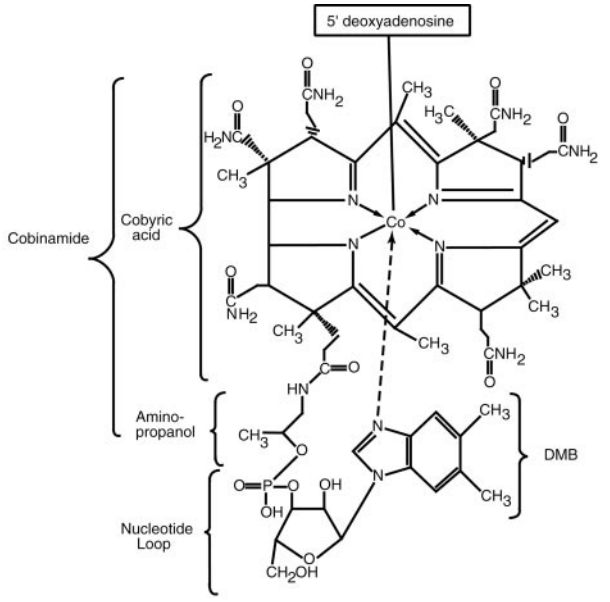


FIG. 1. **Chemical structure of adenosylcobalamin.** In *S. enterica*, the lower ligand is 5,6-dimethylbenzimidazole when grown under aerobic conditions (17).

amide-GDP and ribosylated lower ligand to form the end product, adenosylcobalamin (16). CobT (dimethylbenzimidazole phosphoribosyltransferase) and CobC work sequentially to couple the lower ligand to a ribosyl moiety.

In *S. enterica*, the lower ligand of cobalamin is 5,6-dimethylbenzimidazole when the organism is grown under aerobic conditions (17); however, under anaerobic conditions adenine replaces DMB (18). The transfer reaction in both cases is carried out by CobT (Scheme 1). This enzyme has been shown to be able to phosphoribosylate (*in vitro*) a variety of other bases, including benzimidazole, imidazole, histidine, and guanine (19). There is considerable evidence that broad specificity is a common feature of CobT homologues, since several organisms, including *Escherichia coli*, will incorporate a variety of exogenous bases into functional corrinoids *in vivo* (20).

The three-dimensional structure of CobT from *S. enterica* has been determined complexed with 5,6-dimethylbenzimidazole and its reaction products nicotinic acid and α-ribazole-5′-phosphate (21). This revealed that NaMN:DMB phosphoribosyltransferase is a molecular dimer, where each subunit consists of a large and a small domain (Fig. 3). The large domain exhibits a Rossmann fold that is characterized by a parallel six-stranded β-sheet with connecting α-helices. The small domain is made from components of the N- and C-terminal sections of the polypeptide chain and contains a three-helix bundle. As expected, the active site lies at the C-terminal end of the β-strands of the Rossmann fold, which is responsible for coordinating the ribonucleotide, although the orientation of the substrate is opposite to that seen in other nucleotide dependent enzymes. The small domains from both subunits contribute to each DMB binding site of the dimer to create a hydrophobic pocket, which accounts for the coordination of nonpolar substrates. At the time of the original structure determination, it was unclear how the DMB binding pocket could accommodate polar ligands such as adenine.

The variability in the lower ligands found in natural cobamides raises an interesting question of how these diverse bases are recognized by CobT homologues in these microorganisms, how an organism chooses its lower ligand, and how these variants of cobalamin are utilized by enzymes that are dependent on this cofactor. To provide an insight into these questions, a structural study of the interaction between CobT from *S. enterica* and all of the commonly observed lower ligands has been undertaken. We report here the structures of CobT complexed with adenine, 5-methylbenzimidazole, 5-methoxybenzimidazole, *p*-cresol, and phenol. In addition, adenine, 5-methylbenzimidazole, 5-methoxybenzimidazole, and 2-hydroxypurine were observed to react with NaMN within the crystal lattice and undergo the phosphoribosyl transfer reaction to form product. Interestingly, *p*-cresol and phenol did not react with NaMN. These studies provide an insight into the broad specificity of this enzyme. In addition, they also suggest differences that must be present in CobT homologs where the substrate specificity is radically different from that of *S. enterica*. The wide variability of ligands utilized by CobT suggests that the choice of lower ligand by an organism is dictated by earlier enzymatic or metabolic steps that control the nature of the nitrogenous base available to this phosphoribosyltransferase.

MATERIALS AND METHODS

Protein Purification—CobT was overexpressed and purified as described before (19, 21). All purification procedures and manipulation of the protein were carried out at 4 °C. The purified protein was concentrated in a Centriprep-30 concentrator and dialyzed against 20 mM Tris-HCl, pH 7.5, containing 100 mM NaCl. The protein was flash-frozen by dropping 32-μl aliquots into liquid nitrogen that were then stored at −80 °C.

Crystallization and X-ray Data Collection—Crystals of apo-CobT employed for this 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 1.4 M $NH_4H_2PO_4/(NH_4)_2HPO_4$ at pH 6.0 were mixed and suspended over the precipitant solution at room temperature. Crystals grew spontaneously and achieved sizes of 0.6 × 0.6 × 0.2 mm in 2 weeks. 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.

Prior to preparation of the ligand complexes, the crystals of apo-CobT were transferred to 1.4 M $NH_4H_2PO_4/(NH_4)_2HPO_4$, pH 6.0, in which they were indefinitely stable. All of the lower ligands and NaMN were dissolved in this storage solution. Crystals were first transferred to a solution of a lower ligand, and then after 24 h or more a 100 mM NaMN solution was added to a final concentration of 10 mM to observe if the phosphoribosyl transfer might occur within the crystal lattice. The concentrations of the lower ligands and the length of the soak are listed in Table II.

X-ray data were collected at 5 °C with a Siemens HiStar area detector at a crystal to detector distance of 12 cm. 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 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 (22, 23) and internally scaled with XSCALIBRE.[2] Data collection statistics are given in Table II.

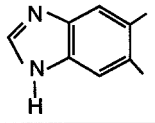
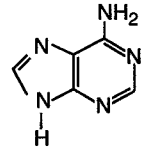
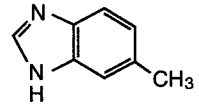
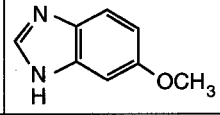
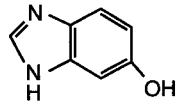
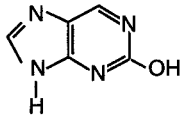
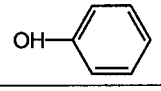
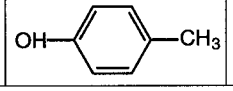
The starting model was obtained by omitting DMB and water molecules from CobT·DMB complex structure (RCSB accession number 1D0S) (21). This initial model was refined by the method of least squares with the program TNT against each data set (25, 26). After this, difference Fourier maps clearly showed the lower ligands or the appropriate products in the active site. These molecules were modeled in the map manually. During the later stages of refinement, solvent molecules were added in locations where the electron density and geometry were consistent with a water molecule utilizing the program PEKPIK in the TNT package (25, 26). The refinement statistics are given in Table III. The coordinates and structure factors have been deposited in the Protein Data Bank (Research Collaboratory for Structural Bioinformatics, Rutgers University) (Table III).

RESULTS AND DISCUSSION

This study of the interaction between CobT and the wide variety of bases that are found as lower ligands in cobamides was facilitated by the ability of crystals of apo-CobT to accept the aromatic bases without destruction of the crystal lattice or induction of a large conformational change in the protein structure. More importantly, the crystal lattice also allows the pro-

[2] G. Wesenberg and I. Rayment, manuscript in preparation.

TABLE I
Lower ligands investigated in this study

Lower ligand	Structure	Microorganisms	Ref.
Dimethylbenzimidazole			
Adenine		*Propionibacterium arabinosum*	(20)
		Propionigenium modestum	(27)
		Clostridium tetanomorphum	(28)
		Methanoplanus limicola	(29)
		All methanococcales tested	(29)
5-methylbenzimidazole		*Desulfobulbus autotrophicum*	(30)
		Desulfobulbus propionicus	(30)
		Archaeglobus fulgidus	(30)
5-methoxybenzimidazole		*Clostridium thermoacetium*	(31)
5-hydroxybenzimidazole		*Pelobacter propionicus*	(27)
		All methanobacteriales tested	(29,32)
		Methanolobus tindarius	(29)
		Methanogenium marisnigri	(29)
		Methanospirillum hungatii	(29)
2-hydroxypurine		This is the analogue of 5-hydroxybenzimidazole	
phenol		*Sporomusa ovata*	(33)
p-cresol		*Sporomusa ovata*	(33)

tein to undergo phosphoribosyl transfer to the base when the second substrate NaMN is diffused into the crystals after the base is bound within the active site. The high resolution data recorded from these complexes are sufficient to identify the orientation of the bound ligand and the stereochemistry of the product. Interestingly, in the absence of the aromatic base, crystals soaked in NaMN always exhibit nicotinate and phosphate within the active site rather than the mononucleotide.

Adenine—Adenine is the lower ligand of cobamides in *Propionibacterium arabinosum*, *Propionigenium modestum*, *Clostridium tetanomorphum*, *Methanoplanus limicola*, and all methanoccales (20, 27–29). It is also the ligand utilized by *S. enterica* when grown under anaerobic conditions (18). Transfer of a crystal of CobT to a solution that contains adenine reveals that this ligand is readily accepted by the active site (Fig. 4*A*) and binds in a very similar position to that of DMB (Fig. 5*A*). As noted previously, the DMB binding pocket is very hydrophobic and is built from components of both subunits (21). This included Met^{177}, Val^{184}, Leu^{315}, and $Pro^{32'}$, $Leu^{30'}$, $Leu^{341'}$, $Ile^{346'}$, and $Leu^{348'}$ from the symmetry-related subunit. Thus, it might be surprising that CobT can utilize a polar substrate such as adenine. The major difference between the binding of adenine and that of DMB in CobT active site is that side-chain oxygen of Ser^{80} hydrogen-bonds to N-10 of adenine and that side-chain amide nitrogen atom of Gln^{88} hydrogen-bonds to the N-3 atom of adenine. In the complex with DMB, Ser^{80} and Gln^{88} are rotated slightly away from DMB to fulfill their hydrogen bonding capabilities, whereas with adenine they move to accommodate its polar nature. Interestingly, the hydrogen-bonding network associated with Gln^{88} would conflict with the alternative orientation of adenine that would place N-10 (of adenine) in close proximity with Gln^{88}.

Previous studies have shown that CobT can utilize adenine *in vitro* (19). To examine whether adenine would react with NaMN within the crystal lattice, as was observed for DMB and NaMN (21), a crystal was transferred first to adenine and then to a solution that contained NaMN. A difference Fourier map calculated at 2.1-Å resolution after one cycle of refinement revealed that the two products (nicotinate and α-adenosine monophosphate) were bound in the CobT active site (Fig. 4*B*). The stereochemistry of the ribonucleotide product is same as that found in the naturally occurring α-adenosine monophos-

phate-containing cobamides, whereby the amino group of adenine is pointing toward the ribose moiety rather than away from it (Fig. 5*B*). As noted above, this arises from the polarity of the DMB binding pocket. The observed orientation of adenine is favored by the hydrogen bond between amide nitrogen atom of Gln^{88} and the N-3 atom of adenine together with the hydrogen bond between the side-chain oxygen of Ser^{80} and N-10 of adenine (Fig. 5*A*). The alternative orientation for adenine would place N-10 of adenine too close to the side chain for Gln^{88}. This interaction excludes the alternative product. This suggests that CobT of *S. enterica* binds adenine with the same stereochemistry as its homologues in the organisms listed above that utilize this base.

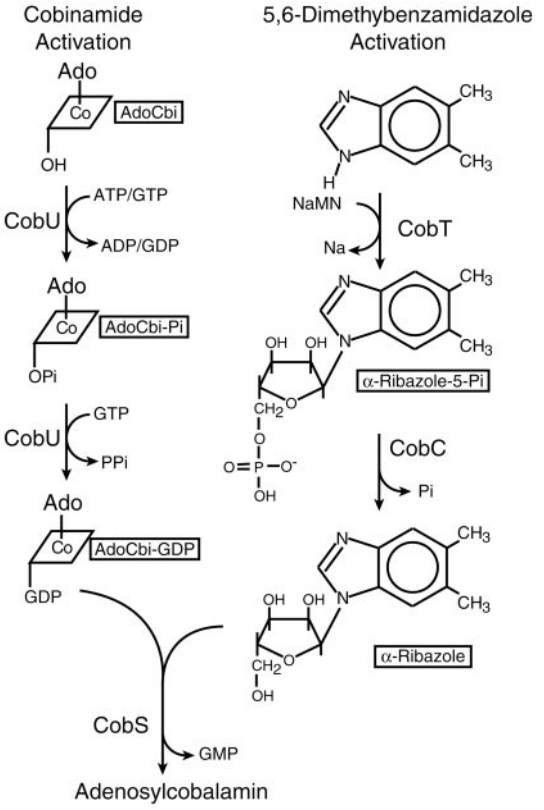
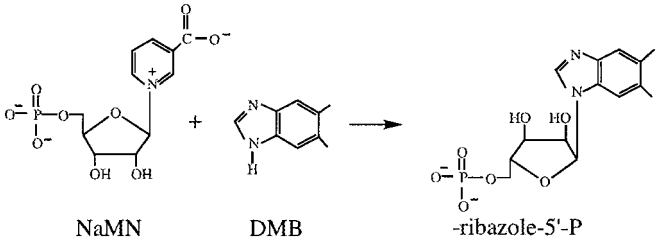


FIG. 2. **A schematic chemical representation of chemical steps involved in nucleotide loop biosynthesis and assembly in adenosylcobalamin biosynthesis in *S. enterica* together with the names of the genes responsible for these transformations.**

5-Methylbenzimidazole—5-Methylbenzimidazole is the lower ligand in microorganisms such as *Desulfobulbus autotrophicum, Desulfobulbus propionicus*, and *Archaeoglobus fulgidus* (30). Crystals soaked in 1.9 mM 5-methylbenzimidazole show unambiguous electron density for this ligand (Fig. 4*C*). The mode of binding is very similar to that of DMB. Transfer of a crystal first to 5-methylbenzimidazole and then to NaMN shows that the phosphoribosyl transfer reaction will also occur within the crystal lattice (Fig. 4*D*). Unequivocal density for N^1-(5-phospho-α-ribosyl)-5-methylbenzimidazole is present within the active site.

In principle, two products could arise from the transfer of methylbenzimidazole to the phosphoribose, yielding either 5- or 6-methylbenzimidazole, depending on the orientation of the benzimidazole moiety in the active site. Significantly, the 5-methyl-substituted product is generated in the crystal lattice as seen in natural cobamides (30). This suggests that the substrate binding pocket has a higher affinity for the methyl group in the 5-position than in the rotated position that would yield the 6-methyl product. The physical basis of this preference is evident from the structure. Although the DMB binding pocket is highly hydrophobic, the inclusion of Ser^{80} at the periphery of the enclosure makes one side of the pocket somewhat more



SCHEME 1

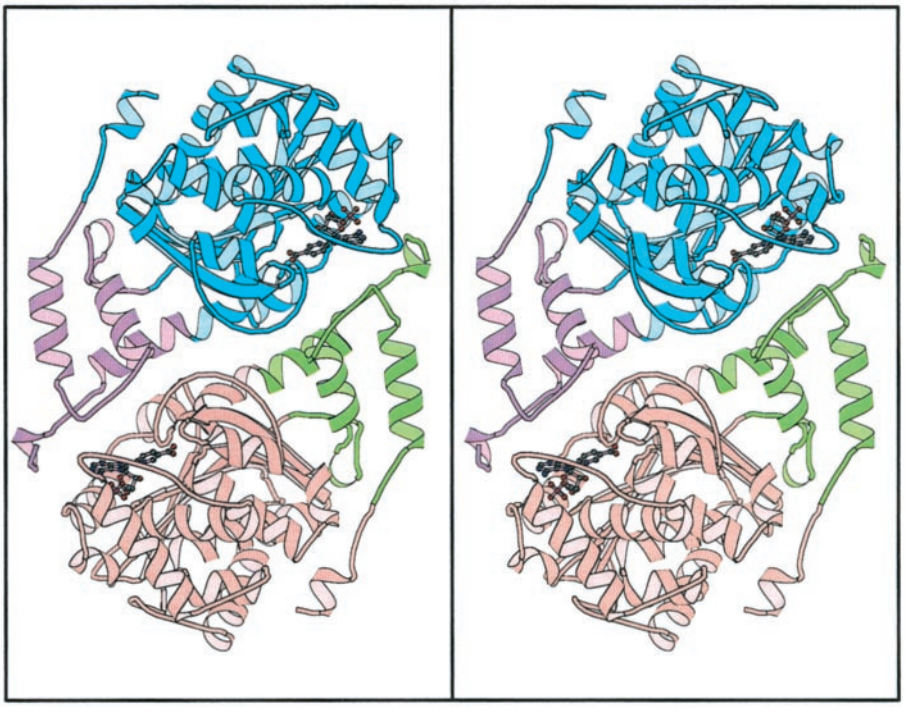


FIG. 3. **Ribbon representation of the CobT dimer complexed with α-adenosylmonophosphate and nicotinate.** CobT is a dimer where each subunit consists of two domains: a large domain dominated by a parallel six-stranded β-sheet with connecting α-helices and a small domain assembled from components of the N- and C-terminal sections of the polypeptide chain. The large domains are depicted in *blue* and *pink*, whereas the small domains are colored in *magenta* and *green*. 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. Figs. 3–7 were prepared with the programs Molscript and Bobscript (43, 44). The figure was prepared from coordinates with Protein Data Bank accession number 1JHA.

TABLE II
Data collection statistics

	Adenine[a]	Adenine NaMN[b]	5-Methylbenzimidazole[c]	5-Methylbenzimidazole NaMN[b]	5-Methoxybenzimidazole[c]
Concentration (mM)	0.6	0.6 10	1.9	1.9 10	13
Length of soak (days)	1	1	3	20	30
Number of crystals	1	1	1	1	1
Resolution (Å)	1.8	2.0	2.2	2.0	2.2
Average I/σ	15.5 (2.7)	11.4 (3.3)	9.5 (2.8)	13.3 (3.8)	11.5 (3.3)
Unique reflections	25,821	20,370	13,880	20,347	14,758
Redundancy	2.9 (1.3)	3.1 (1.5)	1.6 (1.1)	3.9 (1.7)	2.2 (1.3)
Completeness (%)	86.6 (64.4)	89.0 (68.8)	93.9 (85.7)	88.9 (67.8)	94.0 (85.7)
R_{merge}[d]	0.042 (0.141)	0.051 (0.132)	0.049 (0.156)	0.047 (0.118)	0.046 (0.126)

	5-Methoxybenzimidazole NaMN[b]	2-Hydroxypurine NaMN[b]	*p*-Cresol[b]	*p*-Cresol NaMN[b]	Phenol[b]	Phenol NaMN[b]
Concentration (mM)	13 10	10 10	10	10 10	10	10 10
Length of soak (days)	30	1	10	14	20	17
Number of crystals	1	1	1	1	1	1
Resolution (Å)	2.0	2.0	2.0	2.0	2.0	2.0
Average I/σ	12.8 (3.7)	23.1 (7.9)	10.7 (2.7)	13.2 (4.0)	16.0 (4.7)	24.0 (8.7)
Unique reflections	20,523	20,997	20,347	20,349	20,898	20,940
Redundancy	3.2 (1.6)	3.3 (1.6)	3.3 (1.6)	3.3 (1.7)	3.3 (1.6)	3.3 (1.6)
Completeness (%)	88.9 (67.8)	88.8 (67.4)	88.9 (67.6)	88.9 (68.0)	88.8 (67.6)	89.0 (67.7)
R_{merge}[d]	0.046 (0.115)	0.036 (0.063)	0.056 (0.161)	0.054 (0.123)	0.038 (0.098)	0.044 (.061)

[a] The values in parentheses are from resolution shells of 1.85–1.80.
[b] The values in parentheses are from resolution shells of 2.06–2.00.
[c] The values in parentheses are from resolution shells of 2.26–2.20.
[d] $R_{merge} = \Sigma[|Ihi| - |Ih|]/\Sigma Ihi \times 100$, where Ihi and Ih are the intensities of individual and mean structure factors.

polar. As a consequence, the 5-methyl group lies away from the Ser^{80} in such a way that it optimizes the nonpolar interactions. Interestingly alignment of 12 orthologous enzymes from a variety of bacteria (data not shown) reveals that the equivalent residue is mostly serine or tyrosine, although occasionally an alanine. This suggests that a large number of bacteria should be able to incorporate adenine into the position of the lower ligand for cobalamin.

5-Methoxybenzimidazole—5-Methoxybenzimidazole is the lower ligand found in cobamides isolated from *Clostridium thermoaceticum* (31). Crystals soaked in 13 mM 5-methoxybenzimidazole showed unequivocal electron density for 5-methoxybenzimidazole in the active site (Fig. 4*E*). As anticipated from the structure of 5-methylbenzimidazole, the methoxy group points into the 5-methyl binding pocket. Interestingly, there is no apparent accommodation of the lone pair on the methoxy group; however, there is ample room adjacent to the oxygen for transitory water molecules. Transfer of crystals soaked in 5-methoxybenzimidazole to 10 mM NaMN demonstrated that the phosphoribosyltransferase reaction also occurs within the crystal lattice for this substrate (Fig. 4*F*). Unambiguous density for N^1-(5-phospho-α-ribosyl)-5-methoxybenzimidazole is present within the active site. The reaction product has the same orientation of the methoxy group (in the 5- as opposed to the 6-position) as that generated by the CobT homolog present in *C. thermoaceticum* (31).

2-Hydroxypurine—5-Hydroxybenzimidazole is the lower ligand of cobamides in organisms such as *Pelobacter propionicus*, all methanobacteriales tested, *Methanolobus tindarius*, *Methanogenium marisnigri*, and *Methanospirillum hungaritii* (27, 29, 32). Since this lower ligand is not readily available, 2-hydroxypurine was chosen as an analog of this compound. Crystals of CobT were soaked for 24 h in 10 mM 2-hydroxypurine and then transferred to 10 mM NaMN. These crystals show clear density for the reaction product in the active site (Fig. 6*A*). At the initial stages of the refinement, the orientation of the base in the active site was unclear; however, the final omit map after the addition of water molecules and subsequent refinement revealed that the hydroxyl group is oriented toward the 5-methyl binding pocket.

Interestingly, in contrast to all other complexes, residues Ser^{344}, Asn^{345}, Val^{347}, Leu^{348}, and Pro^{349} are disordered in the phosphoribosyl-2-hydroxypurine complex. These residues are located in the C-terminal loop and are involved in binding of DMB. This suggests that the presence of the polar hydroxyl group of 2-hydroxypurine induces disorder into this region of the molecule. As with adenine, the side-chain amide nitrogen atom of Gln^{88} hydrogen-bonds to the N-3 atom of the 2-hydroxypurine. Otherwise, there are few specific interactions with the polar components of the substrate.

The orientation of 2-hydroxypurine is consistent with that found in cobamides that utilize 5-hydroxybenzimidazole as their lower ligands. However, it is clear from the disorder observed in the complex of 2-hydroxypurine that this is not the best ligand for CobT from *S. enterica*. Thus, it is anticipated that the organisms that utilize 5-hydroxybenzimidazole as their lower ligands will include residues in their active sites that specifically distinguish the position of the hydroxyl group. Unfortunately, although the genomic sequences of several methanogens are known, the sequence similarities between CobT and the suggested orthologs in *Methanothermobactor thermautotrophicus* and *Methanococcus jannaschii* are so low that it is difficult to identify the equivalent residues in the active site.[3]

p-Cresol and Phenol—*p*-Cresol and phenol are the lower ligands in cobamides isolated from *Sporomusa ovata* (33) and

[3] Nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase has been assigned to COG2038 (cluster of orthologous groups) 2038 and belongs to group H (proteins involved in coenzyme transport and metabolism) (1, 2). The clustal alignment of >15 proteins that are proposed to be orthologous to CobT can be on the World Wide Web at www.ncbi.nlm.nih.gov/cgi-bin/COG/palox?COG2038. This list includes the sequence for CobT from *A. fulgidus, E. coli*, *M. thermoautotrophicum*, and *M. jannaschii*. There is very limited sequence similarity between CobT from *S. enterica* and the methanogens. Furthermore, the sequence alignment is unedited and thus cannot be taken as proof of the presence or absence of active-site residues.

TABLE III
Least squares refinement statistics (TNT refinement)

	Adenine	Adenine NaMN	5-methylbenzimidazole	5-methylbenzimidazole NaMN	5-methoxybenzimidazole
Resolution limits (Å)	30–1.8	30–2.0	30–2.2	30–2.0	30–2.2
Final *R* factor[a]	17.5	17.8	18.3	17.2	17.9
R_{free}[b]	25.6	27.5	30.2	25.6	29.5
Number of reflections used	25,821	20,370	13,880	20,347	14,758
Number of protein atoms	2420	2420	2420	2420	2420
Number of solvent molecules	118	136	109	150	115
Other molecules, ions	Adenine 2 phosphates	1 AMP 1 nicotinate	5-Methylbenzimidazole 2 phosphates	N^1-(5-phospho-D-ribosyl)-5-methylbenzimidazole 1 nicotinate	5-Methoxy-benzimidazole 2 phosphates
Average *B* values (Å^2)					
Main chain atoms	15.7	19.6	23.9	18.6	21.9
All protein atoms	19.3	22.9	26.6	22.1	25.1
Solvent atoms	29.9	39.9	41.2	39.7	39.7
Weighted r.m.s.[c] deviations from ideality					
Bond lengths (Å)	0.010	0.009	0.010	0.009	0.008
Bond angles (degrees)	1.92	1.78	1.91	1.84	1.78
Planarity (trigonal) (Å)	0.003	0.002	0.003	0.003	0.003
Planarity (others) (Å)	0.011	0.011	0.011	0.011	0.010
Torsional angles (degrees)[d]	16.0	15.6	16.3	15.9	16.2
Protein Data Bank accession number	1JH8	1JHA	1JHM	1JHO	1JHP

	5-methoxybenzimidazo NaMN	2-Hydroxypurine NaMN	*p*-Cresol	*p*-Cresol NaMN	Phenol	Phenol NaMN
Resolution limits (Å)	30–2.0	30–2.0	30–2.0	30–2.0	30–2.0	30–2.0
Final *R* factor[a]	17.8	17.0	18.1	18.1	17.9	17.5
R_{free}[b]	27.9	25.9	28.4	27.4	26.7	26.4
Number of reflections used	20,523	20,997	20,347	20,349	20,898	20,940
Number of protein atoms	2420	2420	2418	2420	2418	2418
Number of solvent molecules	158	150	133	134	141	136
Other molecules, ions	N^1-(5-phospho-D-ribosyl)-5-methoxybenzimidazole 1 nicotinate	N^7-(5-phospho-D-ribosyl)-2-hydroxypurine 1 nicotinate	Cresol 2 phosphates	Cresol 1 nicotinate 1 nicotinate	Phenol 2 phosphates	Phenol 1 nicotinate 1 nicotinate
Average *B* values (Å^2)						
Main chain atoms	15.6	17.3	18.8	17.7	19.9	17.2
All protein atoms	19.0	20.8	22.5	21.2	23.7	20.9
Solvent atoms	36.3	37.0	38.4	38.4	39.2	35.3
Weighted r.m.s.[c] deviations from ideality						
Bond lengths (Å)	0.011	0.010	0.009	0.009	0.009	0.010
Bond angles (degrees)	1.95	1.87	1.86	1.86	1.80	1.88
Planarity (trigonal) (Å)	0.004	0.003	0.003	0.003	0.003	0.003
Planarity (others) (Å)	0.012	0.011	0.011	0.011	0.010	0.011
Torsional angles (degrees)[d]	15.7	16.0	15.9	15.8	15.9	15.7
Protein Data Bank accession number	1JHQ	1JHR	1JHU	1JHV	1JHX	1JHY

[a] $R = \Sigma||F_o| - k \cdot |F_c||/\Sigma|F_o|$.
[b] R_{free} *R* factor for 10% of the data excluded from the refinement. During the final cycles of refinement, all data were included in the refinement. The statistics quoted refer to all of the data.
[c] Root mean square.
[d] No restraints were placed on torsional angles during refinement.

represent the most divergent of all ligands observed in cobamides. In order to examine whether CobT is able to facilitate ribosyl transfer to a phenol moiety, crystals of CobT were soaked for 24 h in 10 mM *p*-cresol or 10 mM phenol. Subsequent data collection and structural analysis revealed that both of these ligands bind in the active site (Fig. 6, *B* and *D*) in approximately the same location as that of DMB (Fig. 7). In these complexes, the oxygen atom of the hydroxyl group of *p*-cresol and phenol form a hydrogen bond to O-ϵ2 of Glu^{317} with a bond-distance of 2.6 and 3.0 Å, respectively.

Crystals soaked in *p*-cresol or phenol were transferred into corresponding solutions that also contained 10 mM NaMN. Difference Fouriers and refinement of the corresponding structures reveal that active sites still contain *p*-cresol or phenol together with nicotinate and phosphate (Fig. 6, *C* and *E*) where the nicotinate arises as an adventitious hydrolysis product of NaMN. This demonstrates that, at least within the crystal lattice, CobT is unable to transfer a phosphoribosyl moiety to a phenolic oxygen. Interestingly, the substrate NaMN is not observed in the crystal lattice when crystals of CobT are soaked in NaMN either in the absence of the second substrate[4] or in the presence of these phenolic substrates. Rather, nicotinate is observed together with a phosphate ion.

In the complexes with these phenolic substrates, the hydroxyl groups of *p*-cresol and phenol remain coordinated to O-ϵ2 of Glu^{317}. As such, this interaction positions the hydroxyl group at a considerable distance from the anticipated position of the C-1 carbon of ribose as deduced by overlaying the structures of α-ribazole-5′-phosphate complex on the structures with phenol or *p*-cresol. Indeed, the distance between the C-1 carbon of the ribose ring of α-ribazole-5′-phosphate and the hydroxyl

[4] C.-G. Cheong and I. Rayment, unpublished results.

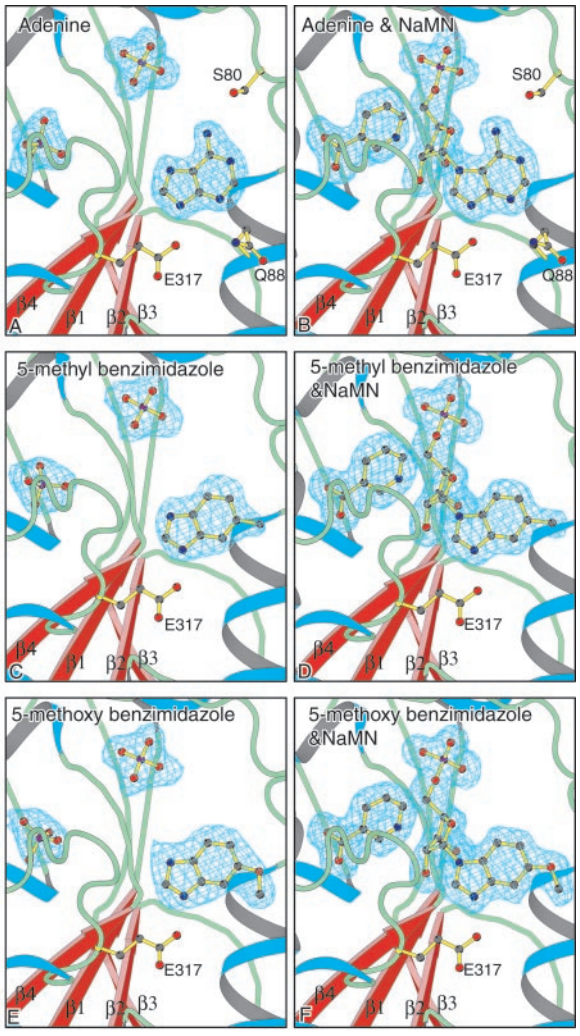


FIG. 4. **Difference electron density for the ligands or products of the reaction with NaMN complexed with CobT.** Shown are adenine (*A*), α-adenosine monophosphate and nicotinate (B), 5-methylbenzimidazole (*C*), N^1-(5-phospho-α-ribosyl)-5-methylbenzimidazole and nicotinate (*D*), 5-methoxybenzimidazole (*E*), and N^1-(5-phospho-α-ribosyl)-5-methoxybenzimidazole and nicotinate (*F*). Coefficients of the form $F_o - F_c$ were utilized where the ligand was excluded from the phase calculation. The maps were contoured at the level of 1σ.

oxygen of *p*-cresol is 4.0 Å when these structures are superimposed. This distance seems to be too long for the reaction to occur. In the complex of phenol and nicotinate, the aromatic ring is somewhat displaced from its position in the binary complex (Fig. 6, *D* and *E*) such that it hydrogen-bonds not only to O-ϵ2 of Glu^{317} but also to N-ϵ2 of Gln^{88} and further displaces the phenolic oxygen from the expected position of the ribose C-1 carbon. This further suggests that the active site of CobT is too large for transfer of these substrates and implies that the CobT homolog in *S. ovata* must have a different active site that serves to bring the two substrates in closer proximity than observed in CobT. In addition, it may require catalytic residues to deprotonate and activate the hydroxyl group of *p*-cresol or phenol that do not appear to be required for phosphoribosyl transfer to an aromatic nitrogenous base.

CONCLUSION

This study provides a structural explanation of why CobT from *S. enterica* will phosphoribosylate all of the aromatic nitrogenous bases commonly found in corrinoids from a very wide range of bacteria and archaea. Furthermore, in all cases, the stereochemistries of the products generated by CobT are identical to those found *in vivo* (Fig. 5*B*). This is surprising, since the polarity of the bases varies considerably from dimethylbenzimidazole to adenine. Examination of the structures reveals that the active site of CobT can accommodate DMB or adenine equally well with only minor changes in the orientation of the side chains to facilitate presentation of a hydrophobic surface to DMB or form compensatory hydrogen bonds to the polar components of adenine. In addition, the binding site contains a distinct hydrophobic polarity that selects for the 5-substituted benzimidazoles in those cases where a choice would otherwise result in the substituent in either the 5- or 6-positions. This arises because the binding site consists of a deep but narrow pocket lined with hydrophobic residues coupled with a hydrophilic lining to the opening to the solvent (Ser^{80}). As a consequence, the aromatic moieties of the products adopt remarkably similar locations in the active site (Fig. 5*B*). These structural observations account for the ability of *S. enterica* to incorporate either DMB or adenine as the lower ligand for cobalamin, depending on the environmental conditions. The ability of CobT to utilize alternative bases as demonstrated in earlier biochemical studies (19) and explained here demands that the choice of lower ligand utilized by an organism is determined by metabolic conditions or as yet undetermined biochemical pathways that generate the nitrogenous bases. Indeed, the broad specificity of the phosphoribosyltransferase activity was noted many years ago when it was shown that several organisms, including *E. coli*, will incorporate a variety of exogenous bases into functional corrinoids *in vivo*. This suggests that broad specificity is a common feature of CobT homologs (20).

Compared with most of the biosynthetic pathway for cobalamin, rather little is known about the synthesis of the nitrogenous base. Genetic searches for mutations that render *S. enterica* unable to make adenosylcobalamin unless provided with DMB all map to the *cobT* gene (34–36). Although it has been established that DMB is most likely synthesized from FMN in the aerobic pathway (37–39), how this is accomplished in molecular detail has not been resolved. Recent work has shown that the C2 carbon of DMB is derived from the ribosyl moiety of FMN (40). Furthermore, there is a 20% preference for N5 of FMN to reside in the N1 position of α-ribazole (41). As a consequence, there has been considerable discussion concerning the potential role of CobT in the biosynthesis of DMB. Indeed, it has been suggested that CobT is a multifunctional enzyme that synthesizes α-ribazole-5′-phosphate from nicotinate mononucleotide and DMB and that also participates in the biosynthesis of DMB (35).

The present study does not address how or whether CobT participates in DMB biosynthesis, but it does demonstrate that the enzyme has great flexibility in its choice of substrate. As a consequence, this structural study identifies one of the reasons why it has been difficult to identify the genes involved in the biosynthesis of the lower ligand. The broad specificity of CobT allows it to incorporate a wide range of inherent and exogenous bases into the lower ligand of cobalamin. Thus, a lesion in one pathway can be readily compensated by a base from an alternative source. CobT functions as the gate keeper for the formation of the lower ligand in that it can accept bases from a wide range of sources.

Finally, the structural studies described here suggest that the enzyme responsible for the phosphoribosyl transfer to phenol or *p*-cresol in *S. ovata* must have an active site that is different from that observed in CobT (33, 42). The inability of CobT to phosphoribosylate these phenolic substrates is readily explained by the additional size and ligand spacing required to coordinate a substituted imidazole, which as a consequence places the phenolic oxygen at least 4 Å from the C-1 of the

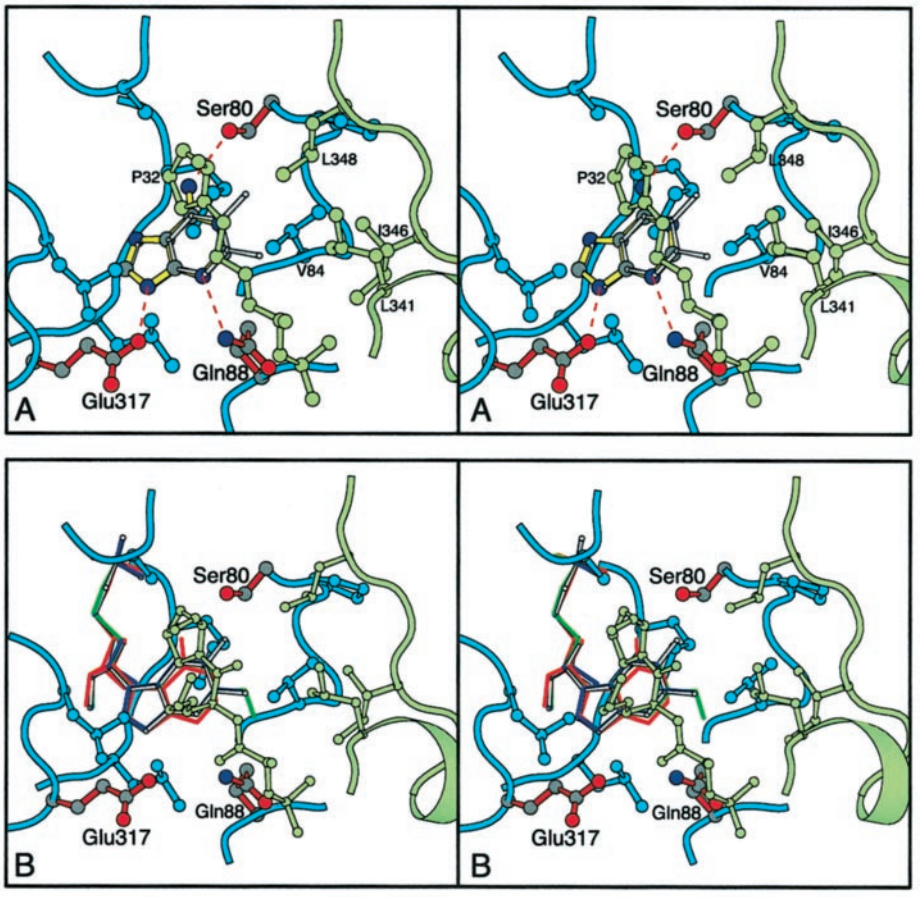


FIG. 5. **Close-up stereoviews of the DMB binding pocket in CobT showing the hydrophobic components that form the binding site.** Side chains and main-chain atoms are depicted in *blue* from the large domain that forms the nicotinate binding site. Symmetry-related interactions from the small domain that constitute many of the hydrophobic interactions are colored in *green*. *A*, the superposition of adenine on DMB. Adenine is shown with *yellow bonds* whereas DMB is depicted with *narrow gray bonds*. The hydrogen-bonding interactions to adenine are shown with *red dashed bonds*. The coordinates for the complex of CobT with DMB correspond to Protein Data Bank accession number 1D0S. *B*, shows phosphoribosyl products for adenine, 5-methylbenzimidazole, and 5-methoxybenzimidazole superimposed on α-ribazole. α-Adenosine monophosphate, N^1-(5-phospho-α-ribosyl)-5-methylbenzimidazole, and N^1-(5-phospho-α-ribosyl)-5-methoxybenzimidazole are depicted as *red*, *blue*, and *green lines*, respectively. This demonstrates that the aromatic moieties of each product adopt very similar locations in the binding pocket, even when they differ considerably in their polarity. The coordinates for the complex with α-ribazole correspond to Protein Data Bank accession number 1D0V. The molecules were superimposed with the program ALIGN (24), modified at the University of Wisconsin by Dr. Gary Wesenberg to allow selection of the segments used for the alignment (available on request).

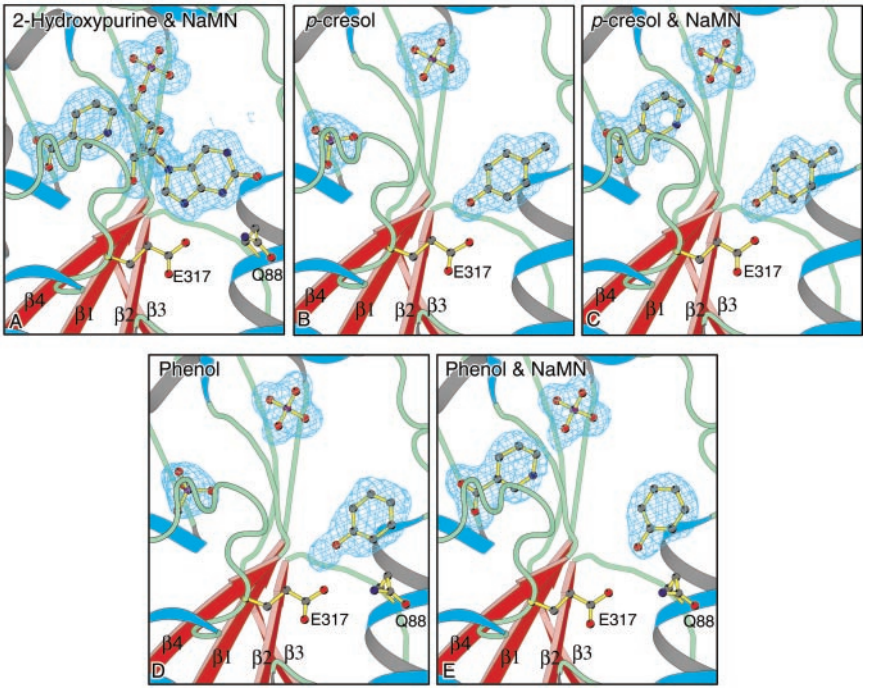


FIG. 6. **Difference electron density for the reaction product for 2-hydroxypurine and NaMN, N^1-(5-phospho-α-ribosyl)-2-hydroxypurine, and nicotinate (*A*), *p*-cresol (*B*), *p*-cresol and nicotinate (*C*), phenol (*D*), and phenol and nicotinate complexed with CobT (E), respectively.** Coefficients of the form $F_o - F_c$ were utilized, where the ligand was excluded from the phase calculation. The maps were contoured at the level of 1σ.

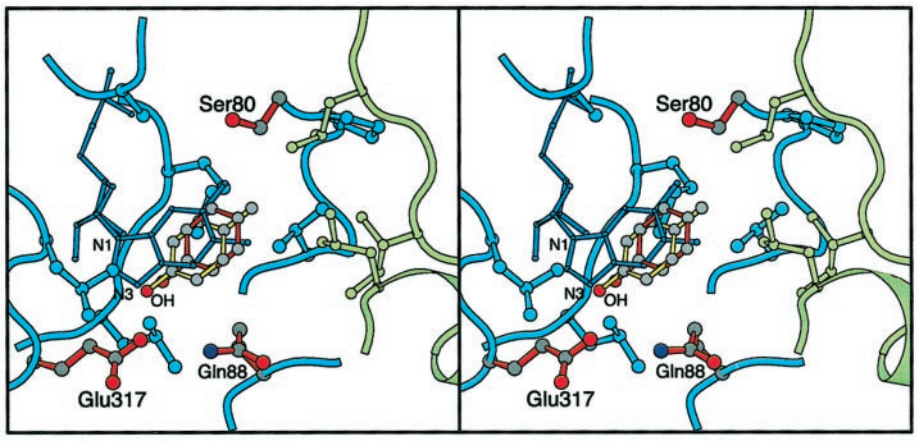


FIG. 7. **Close-up stereoviews of the coordination of phenol and *p*-cresol superimposed on α-ribazole-5′-phosphate in CobT.** Side chains and main-chain atoms are depicted in *cyan* from the large domain that forms the nicotinate binding site. The phenol, *p*-cresol, and α-ribazole-5′-phosphate are depicted with *yellow*, *red*, and *blue bonds*, respectively. Symmetry-related interactions from the small domain that constitute many of the hydrophobic interactions are colored in *green*. The coordinates for the complex with α-ribazole-5′-phosphate correspond to Protein Data Bank accession number 1D0V. The molecules were superimposed with the program ALIGN (24) modified at the University of Wisconsin by Dr. Gary Wesenberg to allow selection of the segments used for the alignment (available on request).

ribose. This suggests that nature has evolved a different enzyme in this organism and suggests that the biosynthetic pathway for cobalamin has yet to yield all of its surprises.

Acknowledgment—We thank Lori Maggio-Hall for helpful discussions.

REFERENCES

1. Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., Kiryutin, B., Galperin, M. Y., Fedorova, N. D., and Koonin, E. V. (2001) *Nucleic Acids Res.* **29,** 22–28
2. Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000) *Nucleic Acids Res.* **28,** 33–36
3. Rondon, M. R., Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1997) *Prog. Nucleic Acid Res. Mol. Biol.* **56,** 347–384
4. Renz, P., Blickle, S., and Friedrich, W. (1987) *Eur. J. Biochem.* **163,** 175–179
5. Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G., and Ludwig, M. L. (1994) *Science* **266,** 1669–1674
6. Mancia, F., Keep, N. H., Nakagawa, A., Leadlay, P. F., McSweeney, S., Rasmussen, B., Bösecke, P., Diat, O., and Evans, P. R. (1996) *Structure* **4,** 339–350
7. Shibata, N., Masuda, J., Tobimatsu, T., Toraya, T., Suto, K., Morimoto, Y., and Yasuoka, N. (1999) *Struct. Fold Des.* **7,** 997–1008
8. Ishida, A., Ichikawa, M., Kobayashi, K., Hitomi, T., Kojima, S., and Toraya, T. (1993) *J. Nutr. Sci. Vitaminol. (Tokyo)* **39,** 115–125
9. Toraya, T., Miyoshi, S., Mori, M., and Wada, K. (1994) *Biochim. Biophys. Acta* **1204,** 169–174
10. 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
11. Scott, A. I. (1993) *Angew. Chem. Int. Ed. Engl.* **32,** 1223–1243
12. Roth, J. R., Lawrence, J. G., and Bobik, T. A. (1996) *Annu. Rev. Microbiol.* **50,** 137–181
13. Thomas, M. G., Thompson, T. B., Rayment, I., and Escalante-Semerena, J. C. (2000) *J. Biol. Chem.* **275,** 27576–27586
14. Thompson, T. B., Thomas, M. G., Escalante-Semerena, J. C., and Rayment, I. (1998) *Biochemistry* **37,** 7686–7695
15. Thompson, T. B., Thomas, M. G., Escalante-Semerena, J. C., and Rayment, I. (1999) *Biochemistry* **38,** 12995–3005
16. Maggio-Hall, L. A., and Escalante-Semerena, J. C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 11798–11803
17. Johnson, M. G., and Escalante-Semerena, J. C. (1992) *J. Biol. Chem.* **267,** 13302–13305
18. Keck, B., and Renz, P. (2000) *Arch. Microbiol.* **173,** 76–77
19. Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1997) *J. Biol. Chem.* **272,** 17662–17667
20. Perlman, D., and Barrett, J. M. (1958) *Can. J. Microbiol.* **4,** 9–15
21. Cheong, C. G., Escalante-Semerena, J. C., and Rayment, I. (1999) *Biochemistry* **38,** 16125–16135
22. Kabsch, W. (1988) *J. Appl. Crystallogr.* **21,** 67–71
23. Kabsch, W. (1988) *J. Appl. Crystallogr.* **21,** 916–924
24. Cohen, G. H. (1997) *J. Appl. Crystallogr.* **30,** 1160–1161
25. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. Sect. A* **43,** 489–501
26. Tronrud, D. E. (1997) *Methods Enzymol.* **277,** 306–319
27. Stupperich, E., and Eisinger, H. J. (1989) *Adv. Space Res.* **9,** 117–125
28. Barker, H. A., Weissbach, H., and Smyth, R. D. (1958) *Proc. Natl. Acad. Sci. U. S. A.* **44,** 1093–1097
29. Stupperich, E., and Krautler, B. (1988) *Arch. Microbiol.* **149,** 268–271
30. Krautler, B., Kohler, H. P., and Stupperich, E. (1988) *Eur. J. Biochem.* **176,** 461–469
31. Irion, E., and Ljungdahl, L. (1965) *Biochemistry* **4,** 2780–2790
32. Krautler, B., Moll, J., and Thauer, R. K. (1987) *Eur. J. Biochem.* **162,** 275–278
33. Stupperich, E., Eisinger, H. J., and Krautler, B. (1989) *Eur. J. Biochem.* **186,** 657–661
34. Escalante-Semerena, J. C., Johnson, M. G., and Roth, J. R. (1992) *J. Bacteriol.* **174,** 24–29
35. Trzebiatowski, J. R., O'Toole, G. A., and Escalante-Semerena, J. C. (1994) *J. Bacteriol.* **176,** 3568–3575
36. Chen, P., Ailion, M., Weyand, N., and Roth, J. (1995) *J. Bacteriol.* **177,** 1461–1469
37. Renz, P. (1970) *FEBS Lett.* **6,** 187–189
38. Renz, P., and Weyhenmeyer, R. (1972) *FEBS Lett.* **22,** 124–126
39. Horig, J. A., and Renz, P. (1980) *Eur. J. Biochem.* **105,** 587–592
40. Keck, B., Munder, M., and Renz, P. (1998) *Arch. Microbiol.* **171,** 66–68
41. Horig, J. A., Renz, P., and Heckmann, G. (1978) *J. Biol. Chem.* **253,** 7410–7414
42. Stupperich, E., Eisinger, H. J., and Krautler, B. (1988) *Eur. J. Biochem.* **172,** 459–464
43. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
44. Esnouf, R. M. (1999) *Acta Crystallogr. Sect. D* **55,** 938–940