Three-Dimensional Structure of Adenosylcobinamide Kinase/Adenosylcobinamide Phosphate Guanylyltransferase (CobU) Complexed with GMP: Evidence for a Substrate-Induced Transferase Active Site[†,‡]

Thomas B. Thompson,[§] Michael G. Thomas,[||] Jorge C. Escalante-Semerena,*[,||] and Ivan Rayment*[,§]

Institute for Enzyme Research and Departments of Biochemistry and Bacteriology, University of Wisconsin, Madison, Wisconsin 53705

Received April 21, 1999; Revised Manuscript Received July 23, 1999

ABSTRACT: The X-ray crystal structure of adenosylcobinamide kinase/adenosylcobinamide phosphate guanylyltransferase (CobU) from *Salmonella typhimurium* bound to GMP has been determined by molecular replacement to 2.2 Å resolution. CobU is a bifunctional enzyme, which catalyzes the phosphorylation of the 1-amino-*O*-2-propanol side chain of the adenosylcobinamide ring and subsequently functions as a guanylyltransferase to form adenosylcobinamide·GDP. The transferase activity involves a covalent enzyme−guanylyl intermediate that is most likely a phosphoramidate linkage to His[46]. Previous studies have shown that the enzyme is a homotrimer and adopts a pinwheel shape. Each subunit consists of a single domain of six parallel β-strands and one antiparallel strand flanked on either side by a total of five α-helices and one helical turn. Interestingly, His[46] in the apoenzyme is located a considerable distance from the kinase active site or P-loop motif and is solvent-exposed [Thompson, T. B., et al. (1998) *Biochemistry 37,* 7686−7695]. To examine the structural relationship of the two active sites, CobU was cocrystallized with GTP and pyrophosphate. Crystals belong to space group $P2_12_12_1$ with the following unit cell dimensions: $a = 58.4$ Å, $b = 87.8$ Å, and $c = 101.6$ Å. The structure shows electron density for the hydrolysis product GMP rather than the expected covalent guanylyl intermediate which appears to have been hydrolyzed in the crystal lattice. Even so, CobU exhibits a substantial conformational rearrangement. The helix axis containing His[46], the site of guanylylation, rotates 30° and translates 11 Å relative to the apo structure and is accompanied by compensatory unwinding and rewinding at the helix ends to allow the induction of a guanosine binding pocket between β-strand 2 and α-helix 2. This conformational change brings the C_α of His[46] approximately 10 Å closer to the P-loop motif such that a phosphate ion located in the P-loop is only 6 Å from the α-phosphate of GMP. This suggests that the P-loop motif may be used to coordinate the terminal phosphates in both the transferase and kinase reactions and implies that the active sites for both reactions overlap.

The complete biosynthesis of cobalamin by bacteria may require up to 30 enzymes (*1*, *2*). The biosynthetic pathway of this exquisitely complex molecule shares a common pathway with heme, siroheme, and chlorophyll synthesis up to the generation of uroporphorinogen III. Thereafter, the pathway to adenosylcobalamin (Figure 1) incorporates many features that are not observed in the biosynthesis of other cyclic tetrapyrroles. There are at least two different pathways for de novo synthesis of the corrin ring. These pathways are called the aerobic and anaerobic pathways (*3*). The key difference between these pathways is the timing of cobalt insertion, which occurs early in the anaerobic pathway (*4*, *5*) and late in the anaerobic pathway (*6*, *7*). Clearly, the anaerobic pathway is less well understood. Even though most of the enzymes and intermediates have been identified, there is still much to be learned about their structures and mechanisms. Currently, the three-dimensional structures of only two enzymes in this pathway are known: CobU and CbiF (*8*, *9*). CbiF, from *Bacillus megaterium*, was determined to 2.4 Å resolution and is one of several transmethylases required for the addition of eight methyl groups to the corrin ring. CbiF contains two α/β-domains and utilizes *S*-adenosyl-L-methionine to methylate cobalt precorrin-4 at the C11 position. On the other hand, CobU is a key participant in the final stages of cobalamin biosynthesis where it is involved in nucleotide loop assembly (*10*).

The final steps of cobalamin synthesis create the amino-propanol side chain of cobalamin and the nucleotide loop that links the lower ligand, 5,6-dimethylbenzimidazole (DMB),[1] to the corrinoid ring. In *Salmonella typhimurium*,

[†] This research was supported in part by NIH Grants AR35186 to I.R. and GM40313 to J.C.E.-S. T.B.T. was supported by NIH Biophysics Training Grant GM08293. BioCARS is supported by NIH Grant RR07707 to Keith Moffat (University of Chicago, Chicago, IL). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Energy Research, under Contract W-31-109-Eng-38.

[‡] The X-ray coordinates have been deposited in the Brookhaven Protein Data Bank under file name 1C9K.

* To whom correspondence should be addressed: Institute for Enzyme Research, 1710 University Ave., Madison, WI 53705. Phone: (608) 262-0529. Fax: (608) 265-2904. E-mail: Ivan_Rayment@biochem.wisc.edu.

[§] Institute for Enzyme Research and Department of Biochemistry.

[||] Department of Bacteriology.

10.1021/bi990910x CCC: $18.00
Published on Web 09/11/1999

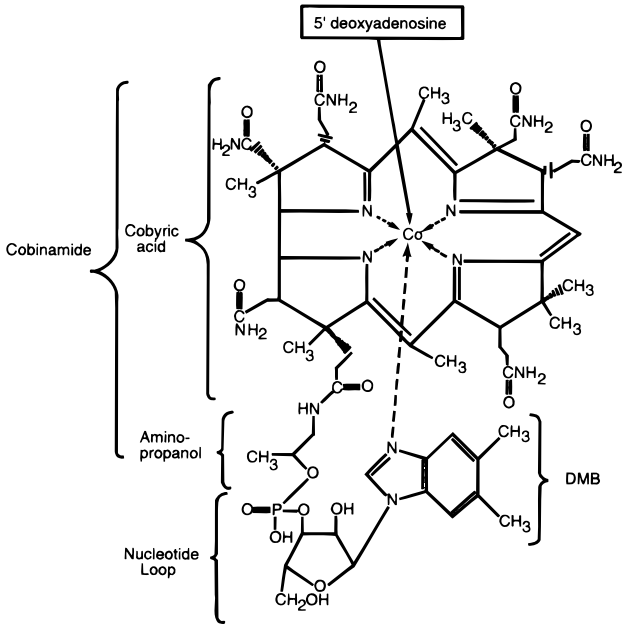

FIGURE 1: Chemical structure of 5′-deoxyadenosylcobalamin.

six enzymes have been identified that catalyze aminopropanol attachment and nucleotide loop assembly: CobU, CobD, CbiB, CobT, CobC, and CobS (*2*, *11*, *12*).[2] These steps can be broken down into four parts: (i) generation of the aminopropanol arm, (ii) activation of adenosylcobinamide, (iii) formation of dimethylbenzimidazole α-riboside or α-ribazole, and (iv) condensation of adenosylcobinamide·GDP (AdoCbi·GDP) and α-ribazole to form adenosylcobalamin (AdoCbl) (Figure 2). Biosynthesis of the aminopropanol arm is mediated by CobD, which is thought to decarboxylate L-threonine phosphate to produce (*R*)-1-amino-*O*-2-propanol phosphate. During de novo biosynthesis, (*R*)-1-amino-*O*-2-propanol phosphate is proposed to be joined to adenosylcobyric acid by the enzyme CbiB to form adenosylcobinamide phosphate. CobU 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. CobT and CobC work sequentially to generate α-ribazole from nicotinic acid mononucleotide and DMB. In the final reaction, CobS combines the activated cobinamide or AdoCbi·GDP and α-ribazole to form the end product, adenosylcobalamin.

CobU or AdoCbi kinase/AdoCbi phosphate guanylyltransferase is a small homotrimeric enzyme (~19 kDa per subunit) that exhibits two distinct enzymatic activities, where it functions both as a kinase and as a nucleotidyl transferase (*10*). In the kinase reaction, CobU catalyzes the phosphorylation of adenosylcobinamide to form adenosylcobinamide phosphate and can utilize a variety of nucleotides as the γ-phosphate donor. Sequence analysis suggested that this activity might be associated with a P-loop located at the N-terminus of the protein (*10*). The second activity adds GMP to adenosylcobinamide phosphate to form adenosylcobinamide·GDP and is specific for GTP. This transferase reaction proceeds through a covalent enzyme−guanylyl intermediate where the putative site for guanylylation is His^{46}. This is the only completely conserved residue that could readily form a phosphoramidate intermediate (*9*, *10*). Thus, this small enzyme is highly unusual in its ability to carry out two very different reactions on a large substrate and implies that there should be two unique active sites on this protein.

Crystallographic studies with the apoprotein revealed that the enzyme is a trimer where the subunits are arranged as a pinwheel with a prominent cleft that might serve as the binding site for adenosylcobinamide (*9*). Each subunit consists of six parallel β-strands and a C-terminal antiparallel β-strand flanked by several α-helices. The fold is similar to several nucleotide-binding proteins which contain a phosphate binding motif or P-loop. Remarkably, CobU shares a similar topology with the central domain of the RecA protein, although they share little or no sequence similarity (*9*). The P-loop motif is most likely the binding site for the nucleotide during the kinase reaction because of its similarity to other kinases that contain this structural motif (*13*). The P-loop is located at the back of a large cleft created from the subunit−subunit interface and was found to be a significant distance from the putative transferase active site.

The structure of the apoenzyme showed that His^{46}, the histidine implicated in the guanylyl transferase activity, was located on the solvent-exposed side of α-helix 2 approximately 20 Å from the kinase active site. From this structure, it was postulated that the aminopropanol arm of the corrin ring was too short to swing between the two distinct nucleotide binding sites and a conformational change was required to bring the kinase and transferase nucleotide binding sites together. Alternatively, it was proposed that the corrin ring might adopt two different conformations when bound to CobU (*9*). This would negate the requirement that the two nucleotide sites be near each other. Clearly, many unanswered questions concerning the relationship between the kinase and transferase sites and the interaction with their respective corrinoid substrates remain.

To investigate the relationship between the two nucleotide binding sites and to examine the specificity for guanosine triphosphate in the transferase reaction, the structure of CobU coordinated to GMP was determined to 2.2 Å resolution. The structure reveals that a conformational change has occurred relative to the apo structure that brings the kinase and transferase sites into close proximity. Interaction of the guanine ring with several critical side chains accounts for the specificity for GTP in the transferase reaction and reveals how this is achieved through formation of a substrate-induced nucleotide binding site. The structure also suggests that a relationship may exist between the kinase and transferase nucleotide binding sites that might account for the remarkable ability of this small enzyme to carry out two distinct chemical reactions. It appears that the ancestral protein was most likely a kinase based on the similarity of the overall topology CobU shares with RecA and other kinases and that the transferase activity may have evolved to utilize the existing phosphate binding motif.

[1] Abbreviations: DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; rms, root-mean-square; AdoCbl, adenosylcobalamin; AdoCbi, adenosylcobinamide; DMB, 5,6-dimethylbenzimidazole.

[2] The names for the gene products involved in these pathways will follow those for *S. typhimurium*; however, many of the proteins were first identified in *Paracoccus denitrificans* but have different genetic names (*1*).

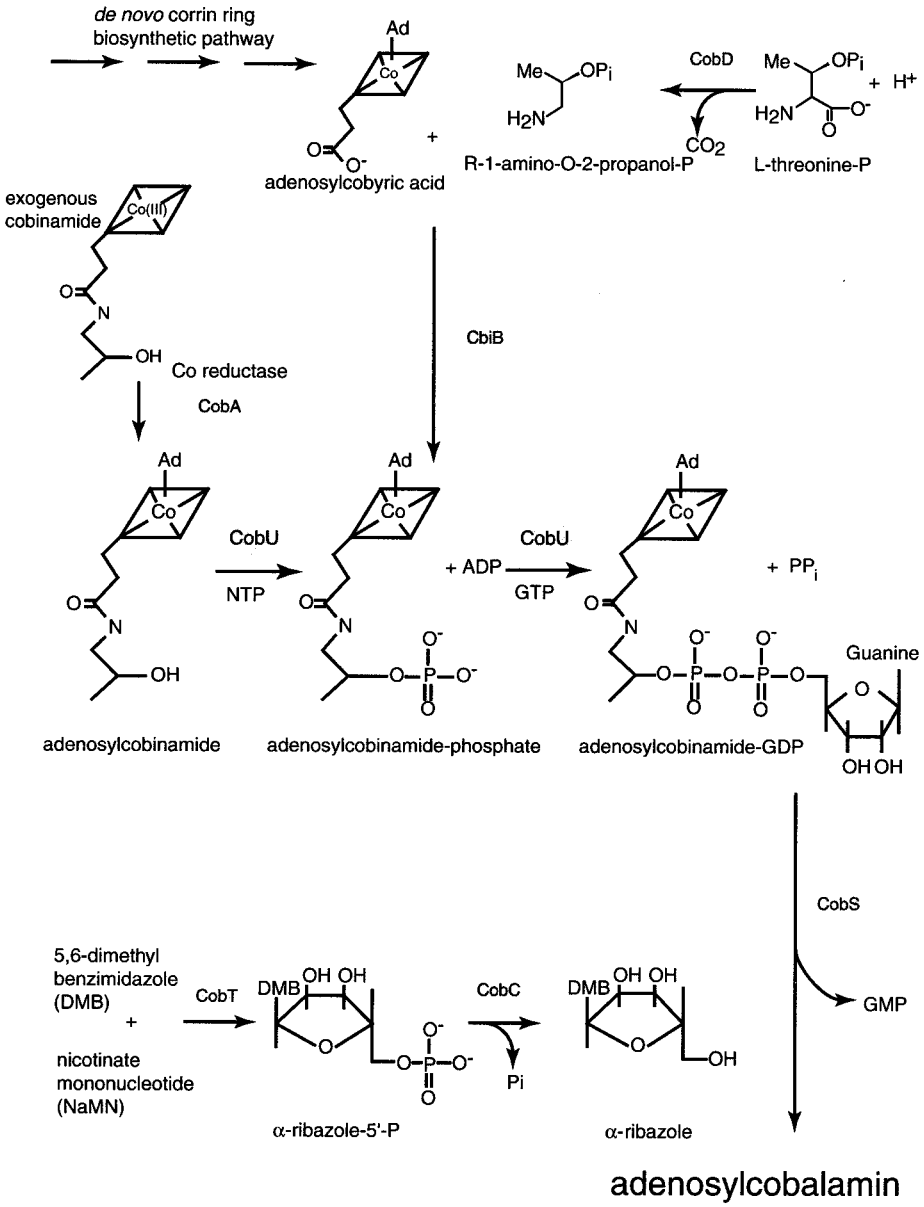

FIGURE 2: Schematic chemical representation of the terminal steps and enzymes (gene names) involved in adenosylcobalamin biosynthesis in *S. typhimurium*.

EXPERIMENTAL PROCEDURES

Crystallization and X-ray Data Collection. CobU was purified as previously described (*9*) and dialyzed against 10 mM Hepes (pH 7.5) and 10 mM DTT at a final concentration of 12 mg/mL. Protein was stored at −80 °C after drop-freezing in liquid nitrogen until required. After being thawed, CobU was preincubated at 4 °C with 2.5 mM pyrophosphate to inhibit the kinase reaction. The protein was then mixed with 5 mM MgGTP and allowed to form the covalent GMP intermediate for a period of 30 min before being mixed with the precipitant.

Crystals were grown by the batch method from 16% (w/v) poly(ethylene glycol) (PEG) 4000, 100 mM $CaCl_2$, and 50 mM CHES (pH 9.0). Optimal conditions for crystal growth required nucleation from prior experiments by streak seeding with a cat whisker (courtesy Harvey Thompson, fat domestic shorthair). Three different crystal morphologies were obtained at slightly different PEG concentrations that belong to at least two different crystal systems. The current study includes X-ray diffraction collected from an orthorhombic crystal which grew to dimensions of 0.4 mm × 0.1 mm × 0.1 mm with space group $P2_12_12_1$ and the following unit cell dimensions: $a = 58.3$ Å, $b = 87.7$ Å, and $c = 101.3$ Å. The crystal lattice contained one trimer in the asymmetric unit. The orthorhombic crystal form was chosen since it was the least sensitive to cryogenic transfer.

For X-ray data collection at −140 °C, crystals were transferred stepwise to a cryoprotectant solution of 25% PEG 4000, 500 mM NaC1, 150 mM $MgCl_2$, 7% ethylene glycol, and 50 mM CHES (pH 9.0). X-ray data were collected to 2.2 Å resolution with the acsd Quantum-1 CCD detector at the Advanced Photon Source of Argonne National Laboratory (Argonne, IL) beamline 14BM-D. Diffraction data were processed with DENZO and internally scaled with SCALEPACK (*14*). Data collection statistics are listed in Table 1.

Structure Determination, Averaging, and Refinement. The structure of the CobU·GMP complex was determined by

Table 1: Data Collection and Refinement Statistics[a]

resolution limits (Å)	30.0−2.20
no. of reflections used	26435
completeness (%)[b]	97.3 (94.9)
average I/σ	27.5 (9.5)
R_{merge} (%)[b,c]	5.3 (22.7)
final *R*-factor[d]	19.1
no. of protein atoms	3820
no. of solvent molecules	336
no. of other molecules or ions	3 GMP, 1 Mg, 1 pyrophosphate, and 2 phosphate
average *B* value	
main chain atoms	37.6
all protein atoms	41.9
solvent atoms	50.1
weighted rms deviations from ideality	
bond lengths (Å)	0.011
bond angles (deg)	2.32
planarity (trigonal) (Å)	0.003
planarity (others) (Å)	0.009
torsional angles (deg)	18.94

[a] TNT refinement. [b] The values in parentheses are from the resolution shells of 2.28−2.00 Å. [c] $R_{merge} = \Sigma[|I_{hi}| - |I_h|]/\Sigma I_{hi} \times 100$, where I_{hi} and I_h are the intensities of individual and mean structure factors, respectively. [d] *R*-factor $= \Sigma||F_o| - k|F_c||/\Sigma|F_o|$.

molecular replacement with the program AmoRe (*15*, *16*). The initial search model included all three subunits of the native structure (*9*). A rotational search of data between 4 and 8 Å resolution yielded three dominant peaks at 27.5, 27.3, and 26.9. A translational search of the three largest peaks yielded a correlation coefficient of 49.3. Rigid-body refinement of this solution resulted in a correlation coefficient of 63.7 and an *R*-factor of 40.8%.

Examination of the molecular replacement model in the crystallographic cell revealed that residues Asp[37]−Ala[52] overlapped with a symmetry equivalent molecule. This was the first indication that a large conformational change had occurred upon nucleotide binding. Due to a lack of electron density in this region and others, residues Ala[18]−Leu[27], Gln[33]−His[53], and Phe[91]−Gln[101] were omitted from all three subunits. This was followed by 10 cycles of least-squares refinement with the program TNT which lowered the *R*-factor from 42 to 36% (*17*). Two further rounds of manual model building and least-squares refinement lowered the *R*-factor to 30%.

Cyclical 3-fold averaging was employed to improve the quality of the electron density (*18*). After each round of model building and least-squares refinement, new matrices were calculated with the program LSQKAB (*19*) for averaging which enhanced the electron density for some of the missing residues (*17*, *20*). The *R*-factor for the final averaged model dropped to 25%, and the final model contained all residues except for the region from residue 33 to 40. The final averaged model also contained guanosine along with a phosphate ion located in the P-loop motif. The averaged model was expanded into the unit cell and subjected to a few rounds of model building and least-squares refinement. Solvent molecules were added using the PEKPIK program of the CCP4 package (*16*) in locations where clear density and geometry consistent with a water molecule were observed. The three crystallographically independent subunits have been arbitrarily designated as subunits A, B, and C for the purpose of discussion and reference to the PDB coordinate file 1C9K. The phosphate of the GMP molecule was evident for two of the subunits, A and B. Figure 3a shows the omit electron density for GMP in subunit A. In the third subunit (C), electron density for guanosine is clearly visible; however, its phosphate moiety is disordered. A phosphate molecule was modeled in the P-loop region in two of the three subunits in a location similar to that observed in the apo structure. Figure 3b shows the $2F_o - F_c$ electron density of the P-loop residues with phosphate bound. Magnesium pyrophosphate instead of phosphate was built into the kinase active site of the third subunit. All electron density maps were generated with SIGMAA weighting (*21*). The current *R*-factor is 19.1% for all measured X-ray data to 2.2 Å resolution with root-mean-square deviations from "ideal" geometry of 0.011 Å for bond lengths, 2.32° for bond angles, and 0.009 Å for groups of atoms expected to be coplanar. Analysis of the backbone dihedral angles with PROCHECK revealed that 90.9% or 408 of the residues conformed to the most favorable regions where the other 9.1% or 41 residues conformed to other allowed regions (*22*). All non-glycine amino acids exhibited favorable dihedral geometry. Refinement statistics are given in Table 1.

RESULTS AND DISCUSSION

CobU was cocrystallized with 5 mM MgGTP and 2.5 mM pyrophosphate. In the presence of MgGTP, CobU forms a covalent enzyme GMP adduct at neutral or basic pH (*10*). Consequently, it was anticipated that the structure would contain a guanylylated intermediate; however, no evidence of the attachment of GMP to His[46] is seen. Instead, the final model contains three GMP molecules with the guanine ring in the anti conformation (the phosphate group is absent in subunit C). Preliminary experiments show that the phosphohistidine bond is not as stable as first believed and has a reasonably short half-life of approximately 21 s at room temperature (M. G. Thomas and J. C. Escalante-Semerena, unpublished results). A phosphate molecule was modeled in the P-loop motif for subunits A and B, whereas Mg• pyrophosphate was modeled in subunit C.

The final model for the molecular trimer contains 512 of the 540 amino acids. Each subunit is composed of six parallel β-strands flanked by six α-helices and a C-terminal antiparallel β-strand. The secondary structural elements are labeled α1−α6 and β1−β7 as shown in Figure 4. There are slight differences between the three crystallographically independent subunits where subunits A, B, and C contain 167, 180, and 165 amino acid residues, respectively, out of 180. Figure 5 shows the relative differences between subunits A, B, and C with and without GMP bound. The rms differences in the positions of the α-carbons when the three independent subunits were superimposed were 0.40, 0.49, and 0.50 Å between subunits A and B, B and C, and A and C, respectively. The rms difference in the positions of the α-carbons of the CobU trimer with and without GMP bound is 3.0 Å with a maximum deviation of 12.7 Å.

Weak density was located in two loop regions. Shortly after β-sheet 2 at residue Ile[34], the electron density becomes weak and reemerges in α-helix 2 at Ala[41]. A second region of weak electron density is observed in the loop connecting α-helix 4 and α-helix 5 (residues Gly[94]−Asn[97]). Both of these

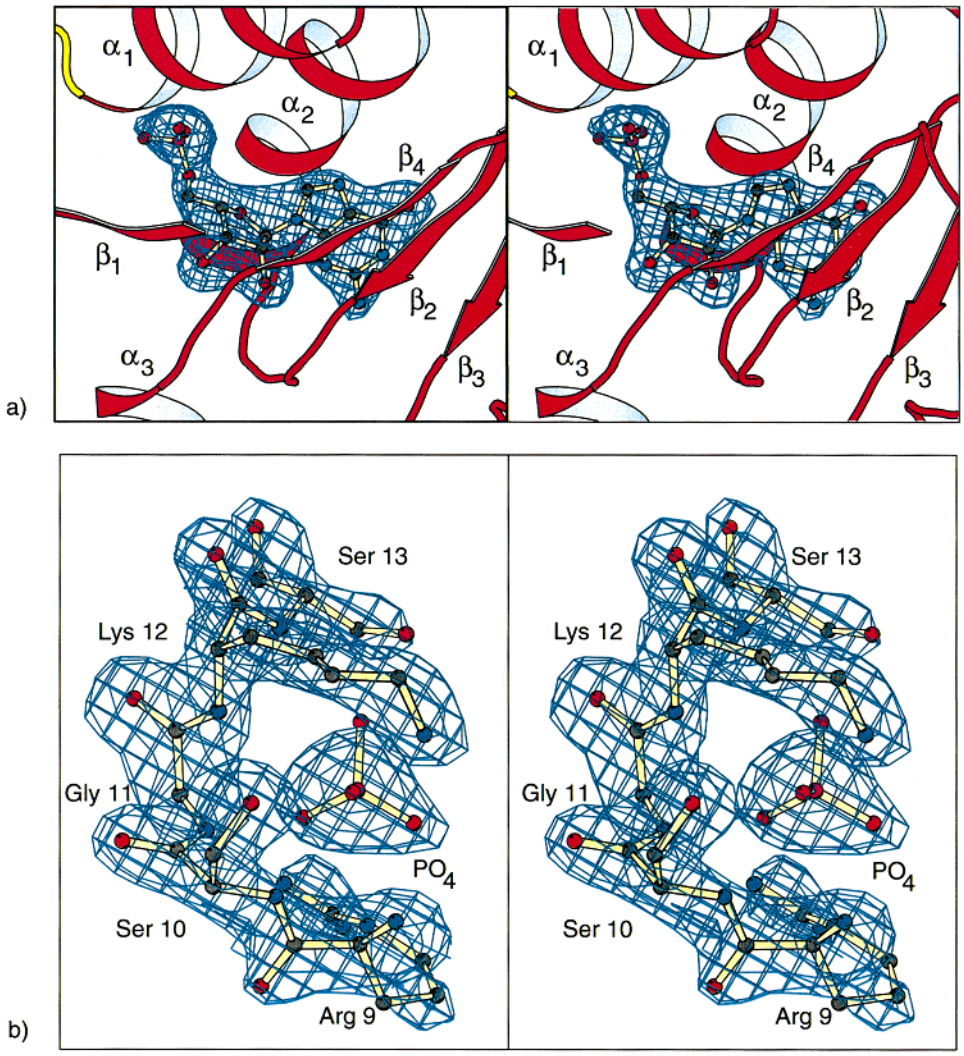

FIGURE 3: Representative electron density. (a) Stereoview of the difference electron density for the guanosine moiety in subunit A. The coefficients for the difference electron density were of the form $F_o - F_c$, where the guanosine molecule was omitted from structural refinement and phase calculation. The figure was prepared with the programs MOLDED and MOLSCRIPT (*36*, *37*). (b) Stereoview of the electron density in the region of the P-loop motif with a bound phosphate. The coefficients for the electron density were of the form $2F_o - F_c$. This figure was prepared with the program BOBSCRIPT.

regions undergo conformational changes relative to that seen in the apo structure. In addition, weak electron density was observed in molecule C near the loop connecting β-sheet 3 and α-helix 3 (residues Glu^{58}−Trp^{60}). Due to conformational flexibility or multiple conformations, it was not possible to model some of these loops even though there appeared to be considerable positive electron density. Molecular averaging did not yield interpretable electron density, suggesting the molecules may adopt slightly different conformations in these areas.

Conformational Change upon Guanylylation. The original goal of this project was to determine the structure of the guanylyl intermediate. As discussed below the phosphoramidate bond is unstable and results in a complex of the enzyme with GMP. The crystal form used to determine the structure can only be obtained from CobU in the presence of GTP under conditions that yield the guanylyl intermediate. This implies that the crystal lattice has trapped the overall conformation of the guanylyl intermediate since the same crystal form cannot be obtained from CobU and GMP. The structure of the enzyme bound to GMP was determined to address several questions concerning the transferase binding site. For instance, what determines the specificity for guanine in the transferase reaction? Also, what is the relationship between the transferase and kinase active sites, if there is one? The apo structure for CobU suggested that a conformational change would be required to bring the two sites together since the aminopropanol arm of adenosylcobinamide was not long enough to reach both active sites by merely swinging between the two active sites.

As seen in the structure presented here, CobU undergoes a significant conformational change when GMP, or more likely GTP, binds. This creates a binding site specific for guanosine that is completely absent from the apo structure. Figure 5 depicts this conformational change in the trimer. The largest movement occurs in α-helix 2. This helix tilts about 30° toward the cleft relative to the apo structure to accommodate the nucleotide. The conformational change also shortens the N-terminal region of β-sheet 2 and significantly lowers α-helix 2 by ~11 Å. This movement of α-helix 2 is responsible for reducing the distance between the P-loop and the putative site for guanylylation and hence the distance between the kinase and transferase active sites. The conformational change also reduces the width of the large cleft

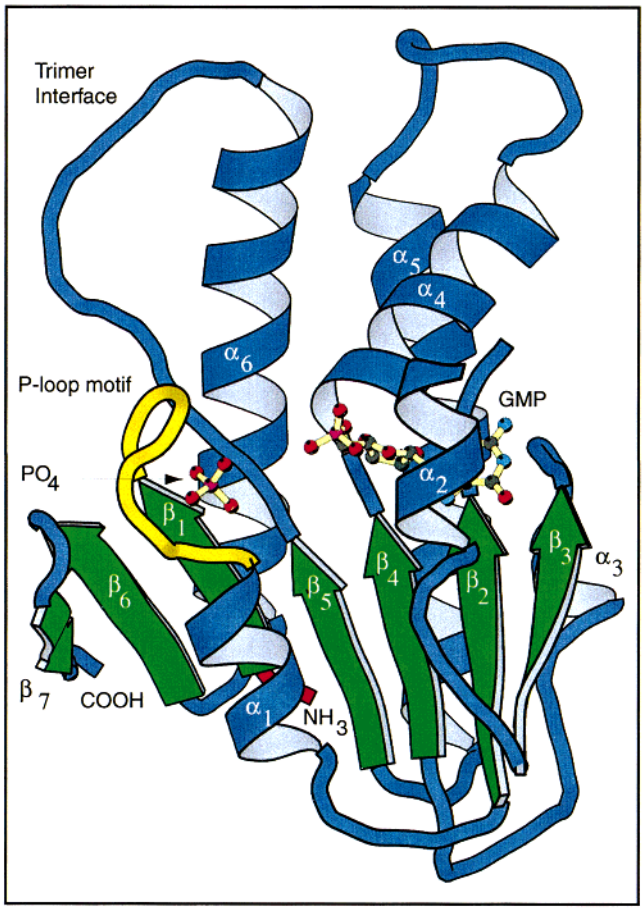

FIGURE 4: Ribbon representation of the CobU monomer with GMP and phosphate bound. This figure was prepared with the program MOLSCRIPT (*36*).

formed from the subunit−subunit interface, thus decreasing its total volume as seen in Figure 5.

Past studies have shown that guanylylation of the enzyme most likely proceeds through a phosphoramidate linkage to His[46] since this is the only conserved amino acid residue, excluding of those found in the P-loop, that could undergo such a modification (*9*, *10*). Recent studies by Thomas and Escalante (unpublished results) have shown that mutation of His[46] removes 95% of the guanylylation. Preparation of the double mutant, H45A/H46A, completely abolishes guanylylation of the enzyme. In the current model, no density is observed for the imidazole ring of His[46]; however, the electron density for the main chain atoms is unequivocal. Furthermore, the residues following His[46] are well-defined which restrict the position of this critical amino acid residue. It would appear that the phosphoramidate bond to His[46] has been hydrolyzed and the structure presented is that of GMP bound after hydrolysis of GTP has occurred. Even with this lack of density, the position of His[46], which lies in α-helix 2, undergoes a large movement as a consequence of binding GTP. When the structures of apo-CobU and CobU·GMP are superimposed, the C_β of His[46] in the GMP complex has translated by two turns of the helix or about 11.4 Å and moves toward the P-loop as shown in the comparison of these structures depicted in Figure 6. In the apo structure, the C_β of His[46] is 15.8 Å from the position of the phosphorus atom of the α-phosphate observed in the GMP complex, whereas in the complex itself, the α-phosphate is only 5.4 Å away. The binding of GTP, as seen in the GMP hydrolysis product described here, induces a conformational change of α-helix 2 that accounts for a movement of 10 Å of the transferase active site toward the P-loop. Even though there is a lack of density for His[46], the position of the *β*-carbon would allow the imidazole side chain to interact with the α-phosphate of GMP with a favorable geometry.

Transferase Binding Site. The transferase reaction is specific for GTP and occurs when GTP and a divalent cation are incubated with CobU, even in the absence of a corrinoid substrate. The current crystal structure exhibits electron density for three guanine nucleotides, two molecules of GMP, and one molecule of guanosine. As shown in Figures 6 and 7, the nucleotide is positioned in the anti conformation so that the α-phosphate is fairly close to the P-loop with the ribose ring and the guanine ring projecting away from the kinase active site. In all subunits, the guanosine base is sandwiched between *β*-strand 2 and the C-terminal end of α-helix 2 where *β*-strand 2 lies toward the distal edge of the large parallel *β*-sheet that extends from the base of the P-loop cleft. The binding of GMP disrupts hydrogen bonding between α-helix 2 and *β*-strand 2 observed in the apo structure and forces α-helix 2 toward the P-loop, partially closing the cleft. The loop connecting *β*-strand 2 to α-helix 2 is largely disordered in all three subunits, especially subunit C where the structure is missing residues Asp[36]−Gln[44]. The variability in the observed order of this loop may be correlated with crystal packing effects; however, the obvious flexibility of this part of the protein may also reflect the structural and energetic constraints imposed by induction of the GMP binding site. Clearly, the region of the protein surrounding His[46] has undergone both a folding and unfolding transition to establish the transferase binding site. In the apo structure, the C-terminal region surrounding His[46] exhibits more conformational flexibility than the remainder of the structure, whereas in the GMP complex, this region is much better defined. From thermodynamic considerations, induction of a well-defined substrate binding site would be entropically unfavorable such that the order−disorder transition observed between *β*-strand 2 and α-helix 2 may help compensate for this effect.

Examination of the environment and hydrogen bonding network surrounding the purine ring reveals why the enzyme preferentially binds guanine nucleotides (Figure 7). As might be expected, there is a specific interaction with the 6-oxo and 2-amino groups of the purine ring that would serve to discriminate between GTP and ATP. There are small differences in the manner in which the crystallographically independent subunits coordinate GMP. Depending on the subunit, the side chain of either Arg[50] or Lys[47] (of α-helix 2) hydrogen bonds to the 6-oxo group of the guanine ring. Arg[50] is completely conserved in all CobU homologues, whereas Lys[47] is either a lysine or arginine. In subunits B and C, only Arg[50] hydrogen bonds to the 6-oxo group of the guanine ring. In all three molecules, the carboxyl side chain of Glu[58] hydrogen bonds with both the 2-amino group and the N1 atom of the guanine ring. Ser[32], which is positioned at the C-terminal end of *β*-strand 2 near the disordered loop, interacts with both the guanine ring and the ribose ring. The hydroxyl side chain of Ser[32] interacts with both the N3 atom of the guanine ring and the 2′-hydroxyl group of the ribose ring. In addition, the amide hydrogen of Ser[32] hydrogen

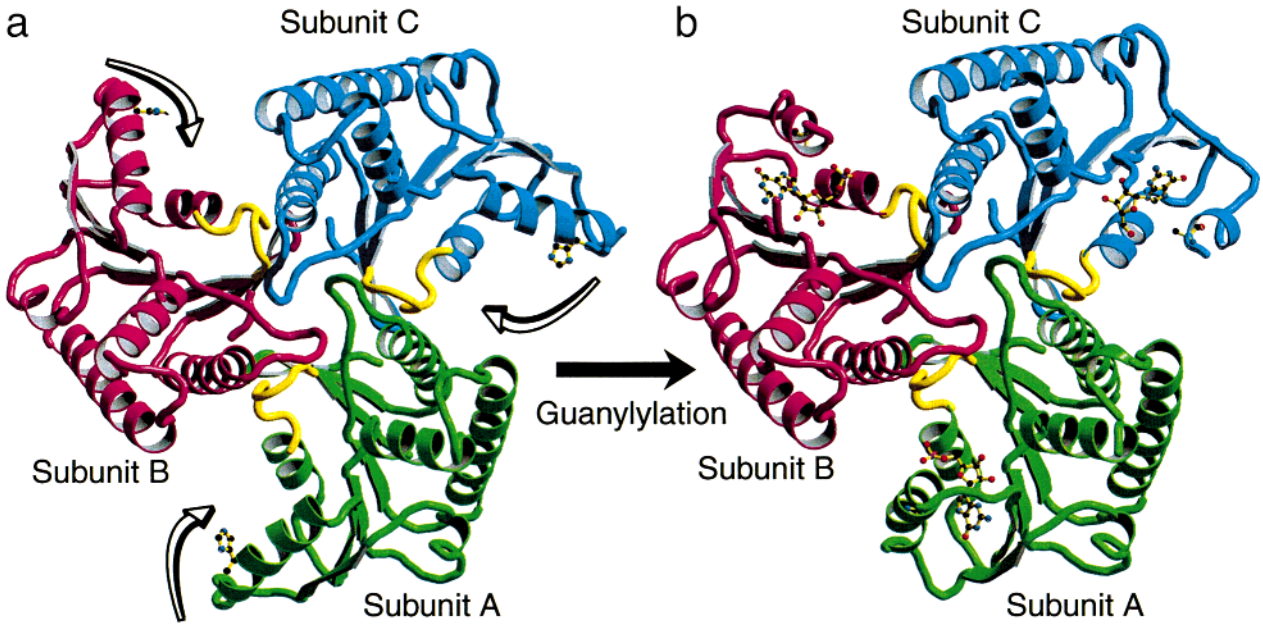

FIGURE 5: Ribbon diagram of the CobU trimer: (a) the apo form of the enzyme and (b) the enzyme bound to GMP. The three subunits are colored separately with the P-loop in yellow and with the putative active site histidine (His^{46}) as a ball-and-stick representation. The arrow in panel a indicates the direction of the large conformational change which occurs upon binding of GTP. The conformational change shrinks the size of the cleft between two adjacent subunits and also brings together the transferase and kinase binding sites of an individual subunit. This figure was prepared with the programs MOLSCRIPT and RASTER3D (*36*, *38*).

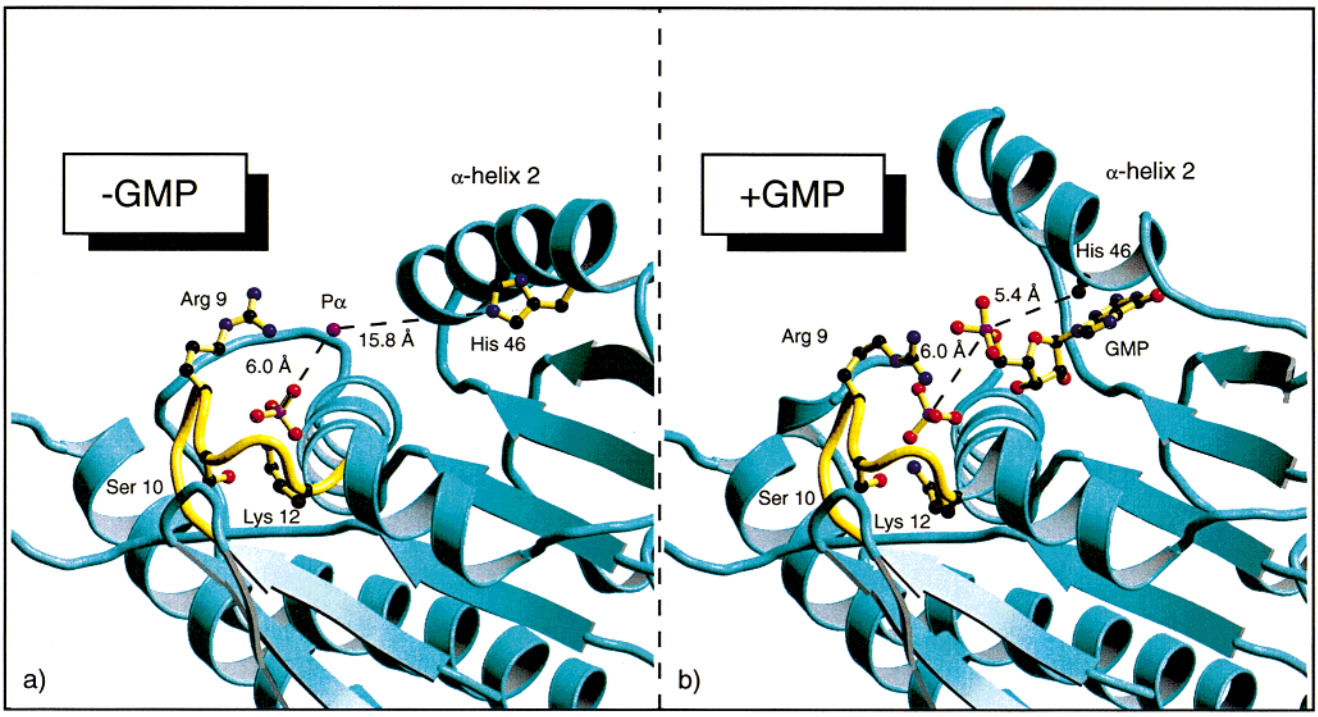

FIGURE 6: Ribbon representation of the helix shift associated with binding GMP where the structure in panel a is that without GMP and the structure in panel b is that of CobU with GMP bound. The figure shows the spatial relationship between the P-loop (yellow) and the transferase active site. Both structures have a phosphate molecule modeled into the P-loop. The position of the guanosine P_α atom in the apo structure was generated by superposition of the GMP structure with the apo structure. The distance from the C_β atom of His^{46} to the P_α atom of the guanosine changes from 15.8 to 5.4 Å due to a shift of α-helix 2 when CobU binds GMP. This figure was prepared with the programs MOLSCRIPT and RASTER3D (*36*, *38*).

bonds to the 2′-hydroxyl group of the ribose ring. The 3′-hydroxyl group of the ribose ring interacts with both the carbonyl backbone of Glu^{80} and the sulfhydryl side chain of Cys^{81}, although it would be difficult to claim a significant hydrogen bonding contribution from this latter interaction. Interestingly, the bond between Glu^{80} and Cys^{81} is a cis peptide bond (as also observed in the apo structure), and its conformation is believed to be critical for both kinase and transferase activity on the basis of preliminary mutational studies (M. G. Thomas and J. C. Escalante-Semerena,

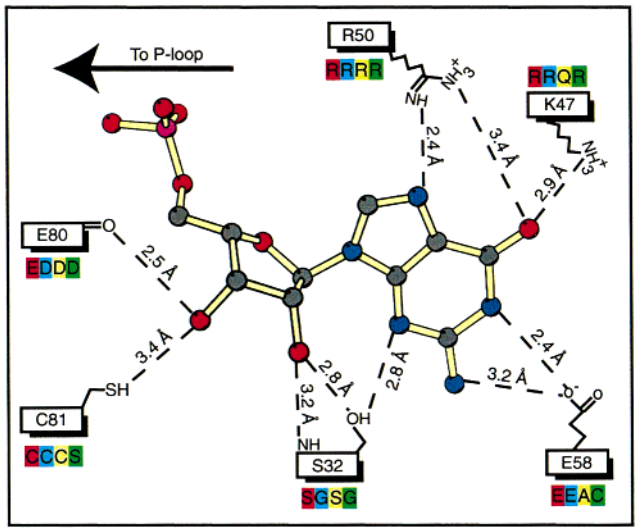

FIGURE 7: Ligand diagram of the amino acids of CobU which interact with the guanine and ribose rings. The letters with the colored background in Figures 7 and 8 indicate the residue type found in CobU homologues for that position. The residues from the species *Escherichia coli*, *Pseudomonas denitrificans*, *Rhodobacter capsulatus*, and *Synechocystis* sp. are depicted in red, blue, yellow, and green, respectively. This figure was prepared with the program MOLSCRIPT (*36*).

preliminary results). The conformations of Glu^{80} and Cys^{81} are essentially identical in both the apo and GMP complex of CobU.

Arg^{50} and Lys^{47} play a key role in the conformational change since they become critical components of the guanine binding site. In the apo structure, both Arg^{50} and Lys^{47} are disordered and solvent-exposed; however, when the enzyme binds guanosine, they hydrogen bond to the guanine ring. Arg^{50} becomes buried in the guanine binding pocket, whereas Lys^{47} remains close to the protein surface. The constellation of interactions with guanine described above and in Figure 7 explains how the enzyme discriminates in favor of GTP. Interestingly, ATP can also modify the enzyme but far less efficiently (M. G. Thomas and J. C. Escalante-Semerena, manuscript in preparation). When CobU is modified by AMP, the enzyme is unable to proceed through the transferase reaction and severely impairs the kinase activity. This suggests that the final orientation of the α-phosphate or the conformation of the nucleotide binding site when AMP is bound might be different. Conversely, it suggests that the specific conformational change brought on by formation of the GMP adduct may be required for the enzyme to transfer GMP to adenosylcobinamide phosphate. Again, it must be noted that CobU is absolutely specific for GTP in the transferase reaction and the reaction with ATP must be considered a dead-end complex; ironically, the CobU·AMP complex is much more stable than the guanylyl intermediate (M. G. Thomas and J. C. Escalante-Semerena, manuscript in preparation).

Pyrophosphate and the P-Loop. A large number of nucleotide-dependent enzymes, including F_1ATPase, myosin, guanylate kinase, adenylate kinase, and RecA, utilize a phosphate binding motif commonly termed the P-loop (*23*). The consensus sequence for this motif is GXXXXGKT/S, and it is often used as a diagnostic tool for identifying nucleotide binding proteins in sequence databases. In most cases, this motif serves to bind the β- and γ-phosphates of a trinucleotide (*13*). In addition, the P-loop usually provides a side chain (Ser or Thr) that functions as a ligand to the divalent cation which usually coordinates the β- and γ-phosphates. The enzymatic activity of these proteins is metal ion-dependent where the role of the metal is both structural and catalytic.

In CobU, there is a conserved P-loop motif located at the N-terminus of the protein with the sequence GGARSGKS. This motif has been implicated in the kinase reaction (*24*). In the CobU·GMP complex, there is a large peak in the electron density located near the residues that make up the P-loop as was observed in the apo structure. In subunits A and B, this peak is consistent with a phosphate ion, whereas in subunit C, density for a pyrophosphate moiety is observed. It was anticipated that magnesium pyrophosphate would be observed in this region since the crystals were grown in the presence of 5 mM magnesium pyrophosphate, a known inhibitor of the kinase reaction (M. G. Thomas and J. C. Escalante-Semerena, unpublished results). Interestingly, magnesium pyrophosphate is observed only in subunit C, whereas phosphate molecules are evident in the other two subunits, although careful examination of the electron density suggests there might be partial occupancy of the magnesium pyrophosphate moiety in these subunits. In all cases, the phosphate moiety common to all subunits (P1) lies in the position observed for the β-phosphate in P-loop enzymes (*13*).

As might be expected from the coordination of the triphosphate moiety in other nucleotide-dependent enzymes in this family (*13*), the amide hydrogens of the P-loop contribute several hydrogen bonds which hold the P1 phosphate in a very strict conformation (Figure 8a). It would appear that the P1 phosphate site has a high affinity because of its presence in all three subunits, including the subunits of the apo structure, especially considering that this anion was not explicitly included in the crystallization solution. In addition to the main chain hydrogen bonds of the P-loop, the side chains of Arg^{9} and Ser^{10}, which are not members of the consensus sequence but are conserved in all CobU homologues, form hydrogen bonds with the P1 phosphate.

The P2 phosphate of pyrophosphate lies in the same location as the γ-phosphate in other nucleotide-dependent enzymes that utilize a P-loop (*13*, *25*), although in this case comparatively few contacts with the P2 phosphate are made. This may account for the lack of pyrophosphate density in two of the three subunits and may be related to the high probability that the P-loop participates in both the kinase and transferase reactions as discussed below (Figure 8b).

The final residue of the P-loop motif is usually a serine or threonine residue that in most cases provides a hydroxyl ligand to the divalent metal ion. The metal ion is evident in subunit C that coordinates a pyrophosphate anion. As indicated in Figure 8a, Ser^{13} interacts with both the metal ion and the P1 phosphate. The hydrogen on the hydroxyl side forms a hydrogen bond with one of the oxygens of the P1 phosphate, whereas the lone pair electrons on Oγ coordinate the magnesium. In addition, the carboxyl group of Glu^{80}, the residue that forms a cis peptide bond with Cys^{81}, contributes to the coordination of the magnesium ion. Interestingly, the main chain carbonyl oxygen is hydrogen bonded to the 2′-hydroxyl of the ribose ring. This interaction

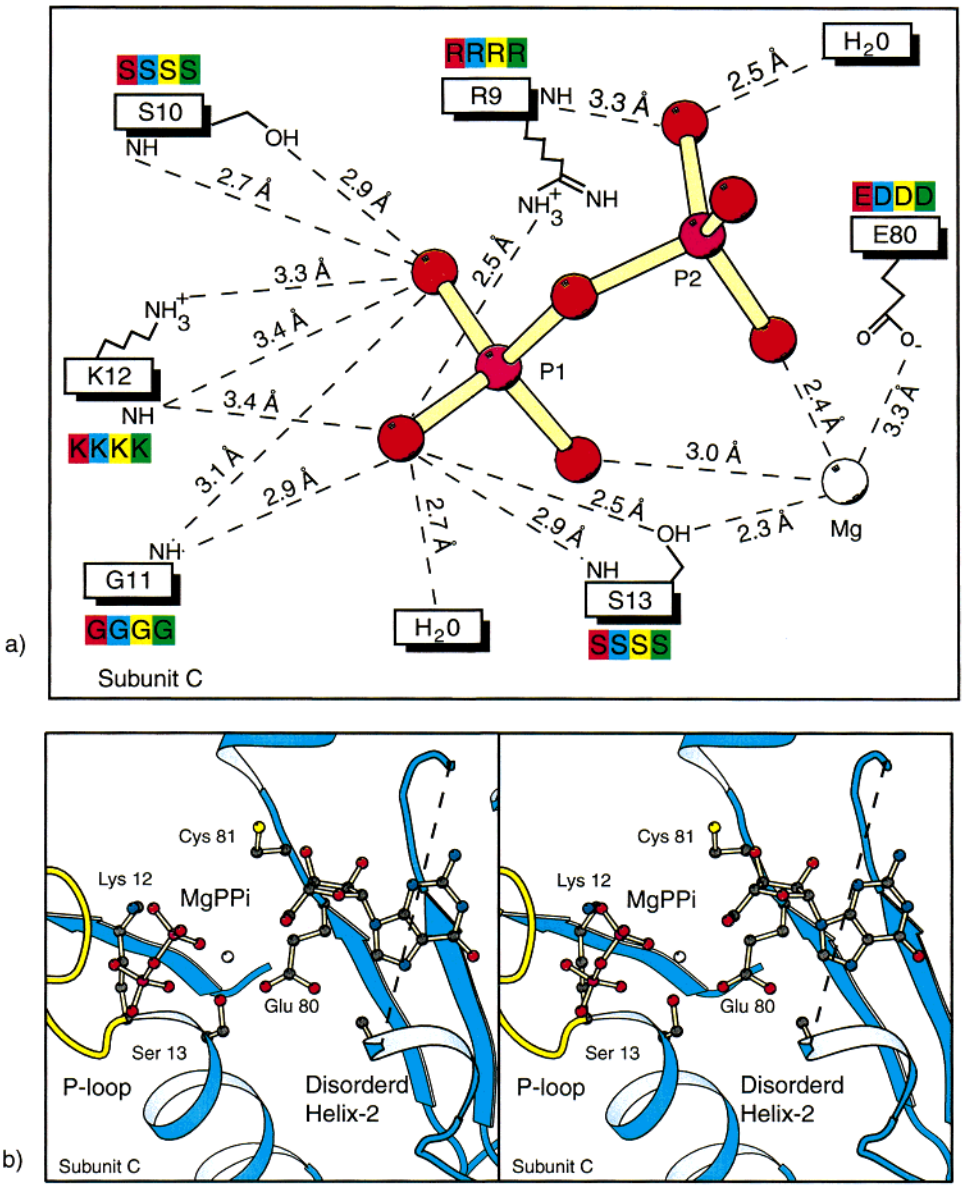

FIGURE 8: Binding of pyrophosphate in subunit C. (a) Ligand diagram of the interactions which are formed when Mg·pyrophosphate binds to CobU. (b) Stereo ribbon diagram displaying the close relationship of GMP to Mg·pyrophosphate. Residues Lys[12] and Ser[13] of the P-loop and Glu[80] and Cys[81] of the cis peptide bond are shown as ball-and-stick representations. The side chains of Glu[80] and Cys[81] are positioned between GMP and Mg·pyrophosphate. This figure was prepared with the program MOLSCRIPT (*36*).

could not be formed if the peptide bond were to assume the normal trans configuration and suggests that the cis peptide bond is important for the catalytic efficiency of the enzyme (Figure 8b). This would be consistent with the general observation that nonprolyl cis peptide bonds are usually only found when they provide a structural or catalytic advantage to the protein (*26*).

The relationship between the pyrophosphate molecule observed in subunit C and the guanine molecule bound in the GMP binding pocket suggests that there is a chemical relationship between the kinase and transferase catalytic sites as shown in Figure 8b. In subunit C, the P2 phosphate of pyrophosphate points toward the guanine molecule and is located approximately 5.4 Å from the 5′-oxygen of the ribose. This suggests that the enzyme may utilize the P-loop to bind the γ-phosphate of GTP during the loading reaction that forms the guanylyl intermediate. Strikingly, the cis peptide bond is positioned between both the kinase and transferase active sites.

CONCLUSIONS

From the structure presented here, it is evident that CobU undergoes a significant rearrangement of α-helix 2 to accommodate binding of a guanine nucleotide. The movement of α-helix 2 translocates His[46] (the putative guanylylation site) ~10 Å closer to the P-loop, significantly reducing the distance between the two active sites. This leads to the hypothesis that the P-loop is utilized to bind the GTP substrate in the first step of the transferase reaction that forms the covalent guanylyl intermediate. From the position of the pyrophosphate in subunit C and the location of the guanine in the GMP binding site, it would appear likely that the γ-phosphate of the GTP may be coordinated by the P-loop during the transferase reaction. The relative position of the GMP binding site and the P-loop suggest that the phosphate

of adenosylcobinamide phosphate should be located somewhere between the phosphate of GMP and the P2 phosphate of the pyrophosphate molecule. Presumably during the kinase reaction, the aminopropanol hydroxyl of adenosylcobinamide will attack the γ-phosphate of ATP (or other nucleotide) as it resides in the P2 position of pyrophosphate, by its analogy to other nucleotide-dependent enzymes. In this position, the phosphate now attached to adenosylcobinamide would be ideally located for attack on the α-phosphate of the guanylyl intermediate to form adenosylcobinamide·GDP. This scenario is consistent with the observation that the transferase activity is inhibited by pyrophosphate (M. G. Thomas and J. C. Escalante-Semerena, preliminary results). The order of substrate binding is unknown at present; however, it seems likely that the type of conformational change observed in the presence of GMP will also be present when adenosylcobinamide binds. Preliminary crystallization experiments suggest the enzyme can bind adenosylcobinamide in the absence of GTP, although it is unknown if this induces the same type of conformational change observed here.

Currently, only three other enzymes besides CobU are classified as guanylyltransferases: (1) mRNA capping enzyme (*27*−*30*), (2) mannose-1-phosphate guanylyltransferase (*31*), and (3) GTP:GTP guanylyltransferase from *Artemia* (*32*). The sequence of CobU is at best only slightly similar to the sequences of these enzymes. As with CobU, both the mRNA capping enzyme and the GTP:GTP guanylyltransferase proceed through an enzyme-bound GMP adduct. A crystallographic analysis of the mRNA capping enzyme revealed that a large conformational change was associated with the formation of the GMP adduct (*30*). Unlike CobU and GTP:GTP guanylyltransferase, the mRNA capping enzyme forms a phospholysine intermediate. Conversely, members of the HIT or "histidine triad" family of proteins such as galactose-1-phosphate uridylyltransferase or the histidine triad nucleotide-binding protein proceed through a phosphoramidate intermediate at a specific histidine residue (*33*, *34*). It is unlikely that CobU falls into the HIT family of proteins since the overall structure and active site architecture are completely different (*35*). Thus, it would appear that CobU utilizes an alternative protein framework to facilitate its guanylyl transferase activity, one that does not require a histidine triad motif.

One of the fascinating issues surrounding CobU is how a comparatively small protein of 180 amino acids facilitates two distinct enzymatic reactions. Indeed, there are very few other enzymes that are able to accomplish such a feat. This presents the question of how such an enzyme evolved. The similarity in the topology of CobU and RecA and the manner in which pyrophosphate is coordinated by the P-loop suggest that CobU was originally a kinase rather than a transferase. The structure of CobU presented here suggests that the GTP binding site for the transferase reaction may have evolved from a conformationally flexible section of the protein to provide an inducible substrate binding site. This order of evolutionary events is consistent with the structure but presents a biochemical problem because it would appear that transferase activity should have evolved before the kinase activity. Examination of the biosynthetic scheme for nucleotide loop assembly (Figure 2) suggests that the kinase reaction serves primarily as a salvage pathway as was proposed elsewhere (*12*). Indeed, de novo cobalamin biosynthesis does not require the kinase activity, whereas the transferase activity is essential for cobalamin biosynthesis. Thus, on the basis of need, it would seem that the transferase activity should have arisen before the kinase activity. Clearly, further biological and structural studies are needed to fully understand how the biosynthetic pathway observed today evolved.

ACKNOWLEDGMENT

We thank Dr. Jim Thoden for help with data collection and Keith Moffat and staff for data collection support at BioCARS.

REFERENCES

1. Blanche, F., Cameron, B., Crouzet, J., Debussche, L., Thibaut, D., Vuilhorgne, M., Leeper, F. J., and Battersby, A. R. (1995) *Angew. Chem., Int. Ed. 34*, 383−411.
2. Rondon, M. R., Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1997) *Prog. Nucleic Acid Res. Mol. Biol. 56*, 347−384.
3. Raux, E., Thermes, C., Heathcote, P., Rambach, A., and Warren, M. J. (1997) *J. Bacteriol. 179*, 3202−3212.
4. Warren, M. J., Bolt, E. L., Roessner, C. A., Scott, A. I., Spencer, J. B., and Woodcock, S. C. (1994) *Biochem. J. 302*, 837−844.
5. Santander, P. J., Roessner, C. A., Stolowich, N. J., Holderman, M. T., and Scott, A. I. (1997) *Chem. Biol. 4*, 659−666.
6. Muller, G., Zipfel, F., Hlineny, K., Savvidis, E., Hertle, R., Traubeberhard, U., Scott, A. I., Williams, H. J., Stolowich, N. J., Santander, P. J., Warren, M. J., Blanche, F., and Thibaut, D. (1991) *J. Am. Chem. Soc. 113*, 9893−9895.
7. Debussche, L., Couder, M., Thibaut, D., Cameron, B., Crouzet, J., and Blanche, F. (1992) *J. Bacteriol. 174*, 7445−7451.
8. Schubert, H. L., Wilson, K. S., Raux, E., Woodcock, S. C., and Warren, M. J. (1998) *Nat. Struct. Biol. 5*, 585−592.
9. Thompson, T. B., Thomas, M. G., Escalante-Semerena, J. C., and Rayment, I. (1998) *Biochemistry 37*, 7686−7695.
10. O'Toole, G. A., and Escalante-Semerena, J. C. (1995) *J. Biol. Chem. 270*, 23560−23569.
11. O'Toole, G. A., Rondon, M. R., Suh, S.-J., Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1996) in *Escherichia coli and Salmonella Cellular and Molecular Biology* (Neidhardt, F. C., Ed.) ASM Press, Washington, DC.
12. Brushaber, K. R., O'Toole, G. A., and Escalante-Semerena, J. C. (1998) *J. Biol. Chem. 273*, 2684−2691.
13. Smith, C. A., and Rayment, I. (1996) *Biophys. J. 70*, 1590−1602.
14. Otwinowski, Z., and Minor, W. (1997) in *Methods in Enzymology* (Carter, C. W. J., Sweet, R. M., Abelson, J. N., and Simon, M. I., Eds.) Vol. 276, pp 307−326, Academic Press, New York.
15. Navaza, J. (1993) *Acta Crystallogr. D49*, 588−591.
16. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
17. Tronrud, D. E. (1997) *Methods Enzymol. 277*, 306−319.
18. Bricogne, G. (1976) *Acta Crystallogr. A32*, 832−847.
19. Kabsch, W. (1976) *Acta Crystallogr. A32*, 922−923.
20. Jones, T. A. (1985) in *Methods in Enzymology* (Wycoff, H. W., Hirs, C. H. W., and Timasheff, S. N., Eds.) Vol. 115, pp 157−171, Academic Press, New York.
21. Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
22. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
23. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) *EMBO J. 1*, 945−951.
24. O'Toole, G. A. (1994) Ph.D. Thesis, University of Wisconsin, Madison, WI.
25. Gulick, A. M., Bauer, C. B., Thoden, J. B., and Rayment, I. (1997) *Biochemistry 36*, 11619−11628.

26. Herzberg, O., and Moult, J. (1991) *Proteins: Struct., Funct., Genet. 11*, 223−229.
27. Martin, S. A., and Moss, B. (1975) *J. Biol. Chem. 250*, 9330−9335.
28. Shuman, S., and Hurwitz, J. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 187−191.
29. Shuman, S., and Schwer, B. (1995) *Mol. Microbiol. 17*, 405−410.
30. Håkansson, K., Doherty, A. J., Shuman, S., and Wigley, D. B. (1997) *Cell 89*, 545−553.
31. Szumilo, T., Drake, R. R., York, J. L., and Elbein, A. D. (1993) *J. Biol. Chem. 268*, 17943−17950.
32. Liu, J. J., and McLennan, A. G. (1994) *J. Biol. Chem. 269*, 11787−11794.
33. Wedekind, J. E., Frey, P. A., and Rayment, I. (1996) *Biochemistry 35*, 11560−11569.
34. Brenner, C., Garrison, P., Gilmour, J., Peisach, D., Ringe, D., Petsko, G. A., and Lowenstein, J. M. (1997) *Nat. Struct. Biol. 4*, 231−238.
35. Lima, C. D., Klein, M. G., and Hendrickson, W. A. (1997) *Science 278*, 286−290.
36. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
37. Fisher, A. J. (1996) *MOLDED*, University of Wisconsin, Madison, WI.
38. Merrit, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869−873.

BI990910X