Three-Dimensional Structure of Adenosylcobinamide Kinase/Adenosylcobinamide Phosphate Guanylyltransferase from *Salmonella typhimurium* Determined to 2.3 Å Resolution[†,‡]

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

ABSTRACT: The X-ray structure of adenosylcobinamide kinase/adenosylcobinamide phosphate guanylyltransferase (CobU) from *Salmonella typhimurium* has been determined to 2.3 Å resolution. This enzyme of subunit molecular weight 19 770 plays a central role in the assembly of the nucleotide loop for adenosylcobalamin where it catalyzes both the phosphorylation of the 1-amino-2-propanol side chain of the corrin ring and the subsequent attachment of GMP to form the product adenosylcobinamide-GDP. The kinase activity is believed to be associated with a P-loop motif, whereas the transferase activity proceeds at a different site on the enzyme via a guanylyl intermediate. The enzyme was crystallized in the space group $C222_1$ with unit cell dimensions of $a = 96.4$ Å, $b = 114.4$ Å, and $c = 106.7$ Å, with three subunits per asymmetric unit. The structure reveals that the enzyme is a molecular trimer and appears somewhat like a propeller with overall molecular dimensions of approximately 64 Å × 77 Å × 131 Å. Each subunit consists of a single domain that is dominated by a seven-stranded mixed β-sheet flanked on either side by a total of five α-helices and one helical turn. Six of the seven β-strands run parallel. The C-terminal strand lies at the edge of the sheet and runs antiparallel to the others. Interestingly, CobU displays a remarkable structural and topological similarity to the central domain of the RecA protein, although the reason for this observation is unclear. The structure contains a P-loop motif located at the base of a prominent cleft formed by the association of two subunits and is most likely the kinase active site. Each subunit of CobU contains a cis peptide bond between Glu[80] and Cys[81] where Glu[80] faces the P-loop and might serve to coordinate the magnesium ion of the triphosphate substrate. Interestingly, His[46], which is the putative site for guanylylation, lies ∼21 Å from the P-loop and is solvent-exposed. This suggests that the enzyme undergoes a conformational change when the substrates bind to bring these two active sites into closer proximity.

Cobalamin represents one of the most complex organic cofactors to be employed in living systems (*1*). Although many aspects of its biosynthetic pathway have been defined, there is still much to be learned about the structure and mechanism of the enzymes involved in this remarkable process. Most of the current information regarding the biosynthesis of cobalamin has been derived from experiments with the bacterial systems, *Pseudomonas denitrificans* and *Salmonella typhimurium* (*2−5*) where the biosynthetic pathway is typically broken down into three components: (i) synthesis of the corrin ring, (ii) attachment of the upper 5′-deoxyadenosine ligand to the cobalt ion, and (iii) synthesis of the lower 5,6-dimethylbenzimidazole ligand and assembly of the nucleotide loop. In *S. typhimurium* four enzymes have been identified in the biosynthesis of the nucleotide loop after the adenosylation step. These are the proteins corresponding to the genes *cobT*, *cobC*, *cobU*, and *cobS*. The genetic and biochemical properties of these enzymes have been partially characterized (*4, 5*).

Nucleotide loop assembly consists of attaching a phosphate, a ribose sugar, and 5,6-dimethylbenzimidazole to the aminopropyl arm of adenosylcobinamide (Figure 1). There are three parts to this process: synthesis of α-ribazole, activation of the adenosylcobinamide, and coupling of adenosyl cobinamide with α-ribazole. Synthesis of α-ribazole is accomplished by the successive action of CobT and CobC, where CobT transfers the phosphoribosyl group of nicotinate mononucleotide to 5,6-dimethylbenzimidazole to yield α-ribazole phosphate (*6*). Thereafter CobC removes the phosphate to produce α-ribazole (*7*). The function of CobU is to activate adenosylcobinamide through the formation of adenosylcobinamide-GDP (*8−10*). α-Ribazole and adenosylcobinamide-GDP are then the proposed substrates for cobalamin synthase (CobS), which catalyzes the final step in the synthesis of adenosylcobalamin (*8−10*).

Of the enzymes involved in nucleotide loop assembly, CobU has been more fully characterized than any other

* To whom correspondence should be addressed. I.R.: Institute for Enzyme Research, 1710 University Ave., Madison, WI 53705. Phone (608) 262-0529; Fax (608) 265-2904; e-mail ivan@enzyme.wisc.edu. J.C.E.-S.: Department of Bacteriology, University of Wisconsin, 1550 Lindon Drive, Madison, WI 53706. Phone (608) 262-7379; e-mail jcescala@facstaff.wisc.edu.

§ Institute for Enzyme Research and Department of Biochemistry.
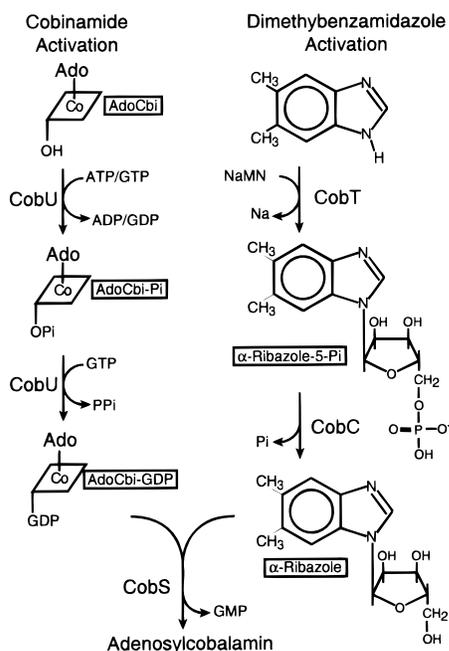
|| Department of Bacteriology.

FIGURE 1: Schematic representation of that part of the cobalamin biosynthetic pathway responsible for nucleotide loop assembly in *S. typhimurium*.

protein in this pathway and has been shown to be an adenosylcobinamide kinase/adenosylcobinamide phosphate guanylyltransferase and has thus two distinct enzymatic activities (*8*, *11*, *10*). The enzyme contains 180 amino acid residues with a total molecular weight of 19 770 (*9*). Despite its small size, this remarkable enzyme catalyzes both the phosphorylation of the 1-amino-2-propanol side chain of adenosylcobinamide and the attachment of GMP to adenosyl-cobinamide phosphate to generate adenosylcobinamide-GDP (*11*, *10*). Interestingly, the kinase activity can utilize either ATP or GTP, whereas the transferase clearly shows a preference for GTP. During *de novo* synthesis of the corrin ring it has been proposed that only the guanylyltransferase activity of CobU is needed for assembly of the nucleotide loop (*12*). Hence, the kinase activity of CobU appears to be needed for the assimilation of unphosphorylated cobina-mide from the environment.

In addition to CobU, there are only three other known guanylyltransferases that have been presently identified: (i) mRNA capping enzyme (*13*−*15*), (ii) mannose-1-phosphate guanylyltransferase (*16*), and (iii) GTP:GTP guanylyltrans-ferase from *Artemia* (*17*). There is no sequence similarity between CobU and either the mRNA capping enzyme or the mannose-1-phosphate guanylyltransferase. Indeed, there is very little sequence similarity between CobU and any other enzyme known at this time except between equivalent proteins in other organisms.

Biochemical data suggest the guanylyltransferase activity of CobU proceeds through a enzyme~GMP intermediate and the site for guanylylation is most likely a histidine (*10*). Since CobU requires GTP for the transferase activity and can utilize either ATP or GTP for the kinase activity, presumably there must be two separate binding sites for these nucleotides and hence two active sites. Thus CobU presents several interest-ing questions, foremost of which is how does such a small enzyme encompass two distinct enzymatic activities? In addition it is unknown how this enzyme coordinates adenos-

ylcobinamide and if there is any relationship between its structure and the cobalamin binding sites in methylmalonyl-CoA mutase (*18*, *19*) and methionine synthase (*20*), whose structures are now known.

A structural study of CobU was initiated to provide answers to the questions outlined above. We report here the structure of the native enzyme to 2.3 Å resolution which shows that the molecule is a molecular trimer and that the putative sites for the kinase and transferase activity are separated by over 21 Å. This information poses further questions about how CobU accommodates both activities.

EXPERIMENTAL PROCEDURES

Protein Purification. The protein was purified in a manner similar to that described before (*10*). Briefly, CobU was overexpressed in strain JE3207. All purification procedures were carried out at 4 °C. Approximately 25 g of cells was resuspended in 0.1 M Tris-HCl, pH 8.0 at 4 °C, containing 10% (v/v) glycerol, 10 mM dithiothreitol (DTT),[1] 1 mM ethylenediaminetetracactetic acid (EDTA), and 1 mM phen-ylmethanesulfonyl fluoride and disrupted by sonication. Cellular debris was removed by centrifugation at $43000g$ for 1 h. Finely ground Ultrapure ammonium sulfate was added to the cell-free extract to 9.6% saturation and clarified by centrifugation. The resulting supernatant was loaded onto a 35 mL bed volume phenyl-Sepharose CL-4B column (Sigma) equilibrated with 0.1 M Tris-HCl, pH 8.0, containing 10% (v/v) glycerol, 10 mM DTT, and 9.6% (saturation) ammonium sulfate. The column was washed with 2 bed volumes of the equilibration buffer. CobU was eluted with 0.1 M Tris-HCl, pH 8.0, containing 10 mM DTT and stored overnight under oxygen-free N_2. Thereafter dye ligand chromatography was used to remove impurities from the protein. $MgCl_2$ was added to a final concentration of 5 mM, and the fractions containing CobU were passed through a 20 mL bed volume column of Cibacron Blue type 3GA (Sigma) equilibrated with 0.1 M Tris-HCl, pH 8.0, containing 10 mM DTT and 5 mM $MgCl_2$. CobU was recovered from the flowthrough and the first bed volume of the wash buffer (0.1 M Tris-HCl, pH 8.0, containing 10 mM DTT and 5 mM $MgCl_2$). The resultant protein was analyzed by SDS−PAGE stained with Coomassie Blue and was estimated to be 90% pure as judged by densitometry. The protein was concen-trated in a Centriprep-10 and dialyzed against 10 mM HEPES, pH 7.5, containing 100 mM NaCl and 10 mM DTT. Typically 150 mg of protein was obtained from ∼25 g of cells.

Crystallization and X-ray Data Collection. Crystals employed for the structural investigation of the native form of CobU were grown by microbatch from 6% poly(ethylene glycol) 3350, 200 mM NaCl, and 50 mM succinate, pH 5.5, at 4 °C at a protein concentration of 5−8 mg/mL. Crystals grew spontaneously or were microseeded and achieved sizes of 0.8 mm × 0.8 mm × 0.3 mm in 14−21 days. Precession photography revealed that the crystals belonged to the space group $C222_1$ with unit cell dimensions of $a = 96.4$ Å, $b = 114.4$ Å, and $c = 106.7$ Å. The crystal lattice contains one trimer per asymmetric unit with a solvent content of 50%. Diffraction maxima were observed to a resolution of 2.1 Å.

Table 1: Data Collection Statistics and Heavy Atom Refinement Statistics

	native[a]	Au(CN)$_2$[a]	ReODADSATP[b]	IrCl$_6$[b]
concn (mM)		0.5	5.0	2.5
length of soak (days)		1	1	3
number of crystals	2	2	1	2
resolution (Å)	2.3	2.5	2.7	2.7
average I/σ	16.0 (1.7)	15.4 (1.7)	6.9 (1.4)	13.0 (1.7)
unique reflections	24 136	19 577	13 267	13 163
redundancy	3.0 (1.9)	4.3 (2.7)	1.5 (1.3)	4.1 (3.1)
completeness (%)	88.5 (64.4)	91.5 (63.9)	87.1 (68.4)	96.9 (90.3)
R_{merge}[c]	4.3 (21.0)	4.6 (29.5)	7.9 (29.3)	5.5
R_{iso}[d]		13.3	34.8	18.4
number of sites		3	3	3
R_{Cullis}[e]		64	66	64
phasing power[f]		1.07 (0.86)	0.91 (1.06)	1.24 (1.13)
figure of merit	0.57 (0.42)			

[a,b] The values in parentheses are from resolution shells of 2.37−2.30 and 2.78−2.70 Å, respectively. [c] $R_{\mathrm{merge}} = \sum[|I_{hi}| - |I_h|]/\sum_{hi}I_{hi}(100)$, where I_{hi} and I_h are the intensities of individual and mean structure factors. [d] $R_{\mathrm{iso}} = \sum[|F_{\mathrm{h}}| - |F_n|]/\sum_n F_n(100)$, where F_{h} and F_n are the heavy atom and native structure factors. [e] $R_{\mathrm{Cullis}} = \sum[|F_{\mathrm{h}} - F_n| - |F_{\mathrm{hc}}|]/\sum|F_{\mathrm{h}} - F_n|(100)$, where the summation is carried out over the centric data only. F_{h} and F_n are the observed heavy atom derivative and native structure factors, and F_{hc} is the calculated structure for the heavy atoms alone. [f] The phasing power is defined as the mean value of the heavy atom structure factor divided by the lack-of-closure error.

For X-ray data collection and preparation of heavy atom derivatives, the crystals were transferred to a synthetic mother liquor consisting of 12% PEG 3350, 400 mM sodium chloride and 50 mM succinate at pH 5.5. Crystals were stable for at least 2 weeks in the synthetic mother liquor, but were extremely sensitive to pH such that all attempts to increase the pH were unsuccessful.

X-ray data were collected to 2.3 Å resolution at −10 °C with a Siemens HiStar area detector at a crystal to detector distance of 14 cm. Cu Kα radiation was generated by a Rigaku RU200 X-ray generator operated at 50 kV and 90 mA and equipped with Siemens Göbel focusing optics. Diffraction data frames of width 0.15° were recorded for 60−90 s. The frames were processed with XDS (*21, 22*) and internally scaled with XCALIBRE (*23*). Table 1 displays the diffraction data statistics for the native and heavy atom derivative data.

Structural Determination. The three-dimensional structure of CobU was determined by the technique of multiple isomorphous replacement with three heavy atom derivatives prepared by soaking native crystals in solutions of 0.5 mM Au(CN)$_2$, 2.5 mM K$_2$IrCl$_6$, or 5.0 mM 2′,3′-ReODADSATP. The rhenium-ATP derivative was synthesized by Ralph Yount, Washington State University, where the rhenium ion is chelated by linker attached to either the 2′ or 3′ oxygen of the ribose. The heavy atom binding sites were located from difference Patterson maps and placed on a common origin by appropriate difference Fourier maps. The statistics for the derivative data collection and heavy atom refinement are given in Table 1. The correct hand of the heavy atom constellation was chosen on the basis of the sign of the anomalous signal in the heavy atom derivatives. The occupancies and positions of all heavy atom sites were refined with the program HEAVY (*24*). Relevant phase calculation statistics are located in Table 1.

Visual inspection of the heavy atom positions for each derivative suggested that they were related by a local 3-fold axis of symmetry that was also evident in preliminary electron density maps. This revealed that the quaternary structure of CobU was trimeric. The transformation matrixes relating the three subunits to each other were derived with the program MUNCHKINS, which examines the correlation

Table 2: Least-Squares Refinement Statistics[a]

resolution limits (Å)	20.0−2.30
Final R-factor[b]	19.8
number of reflections used	26 109
number of protein atoms	4117
number of solvent molecules	208
other molecules, ions	3 phosphate ions
average B values (Å^2)	
main-chain atoms	42.8
all protein atoms	46.9
solvent atoms	55.9
weighted *rms* deviations from ideality	
bond lengths (Å)	0.005
bond angles (deg)	2.07
planarity (trigonal) (Å)	0.004
planarity (others) (Å)	0.005
torsional angles[c] (deg)	18.5

[a] TNT refinement. [b] $R = \sum||F_{\mathrm{o}}| - k|F_{\mathrm{c}}||/\sum|F_{\mathrm{o}}|$. [c] No restraints were placed on torsional angles during refinement.

coefficient between the symmetry-related electron density (*25*). The local 3-fold rotational axis was used to improve the quality of the subsequent electron density map by cyclical molecular averaging with the algorithm of Bricogne (*26*). A self-rotation function was not utilized to detect the orientation of the local 3-fold axis since it was believed at the start of the investigation that the protein was dimeric.

On the basis of the "averaged" electron density map calculated to 2.7 Å resolution, a molecular model for one subunit of the trimer was built. This model was expanded into the unit cell and subjected to numerous cycles of manual model building with the software FRODO (*27*) and least-squares refinement with the program TNT (*28*). During the early stages of the model building the heavy atom phases were combined with model phases with SIGMAA weighting (*29*). The current R-factor is 19.8% for all measured X-ray data from 30 to 2.3 Å with root-mean-square deviations from "ideal" geometry of 0.015 Å for bond lengths, 2.1° for bond angles, and 0.004 Å for groups of atoms expected to be coplanar. Refinement statistics are given in Table 2. A representative portion of the electron density map is shown in Figure 2a.
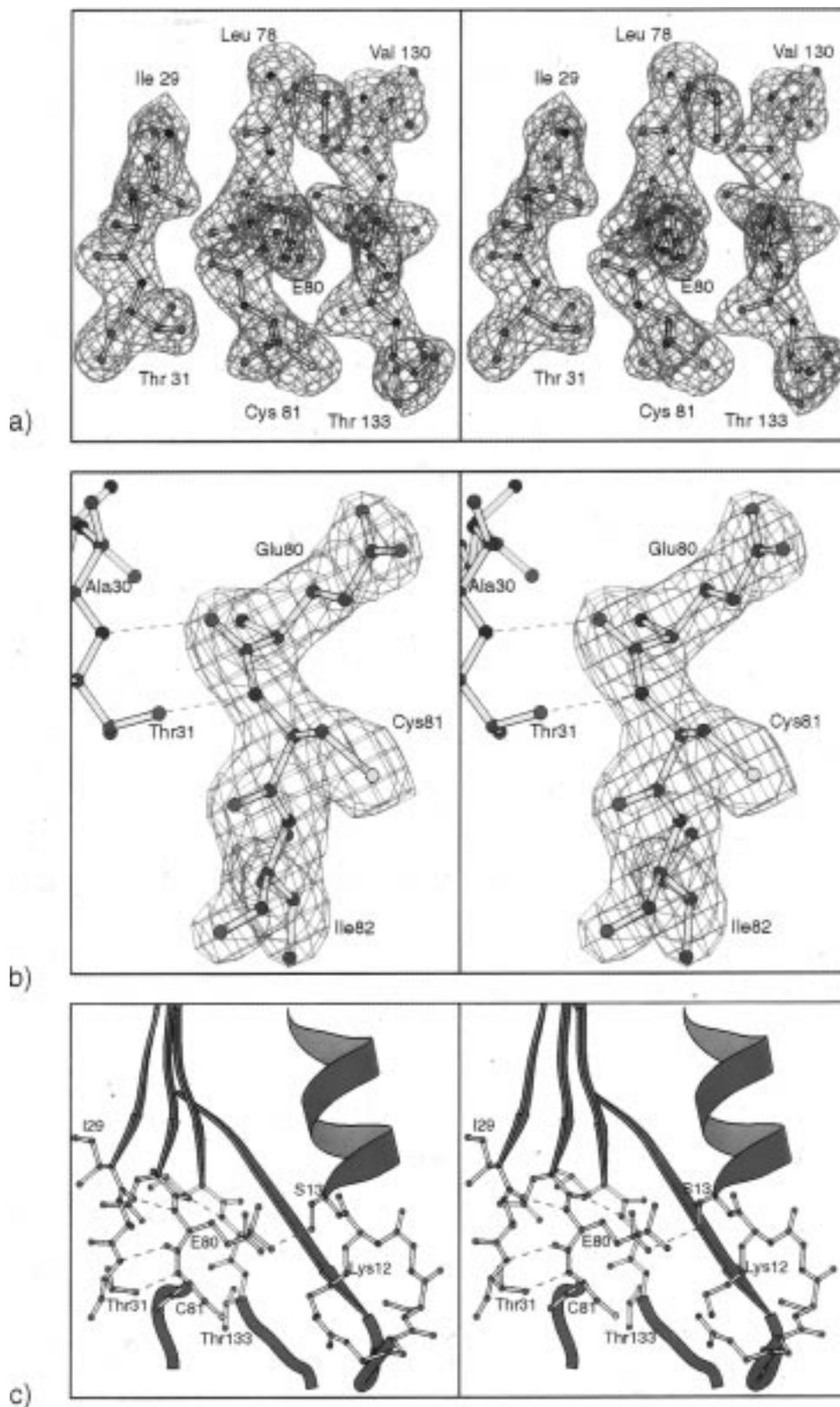
FIGURE 2: Stereoview of (a) a representative portion of the electron density that includes the cis peptide bond between residues Glu[80] and Cys[81] in subunit 1 where the map was contoured at approximately 1σ and was calculated with coefficients of the form $2F_o - F_c$, (b) the difference electron density for the cis peptide bond calculated from coefficients of the form $F_o - F_c$, where Glu80−Ile82 were omitted from the refinement and phase calculation, and (c) an expanded view of the hydrogen-bonding pattern surrounding this cis peptide bond. This figure was prepared with the programs MOLDED and MOLSCRIPT (*58, 59*).

RESULTS AND DISCUSSION

The final model for the CobU trimer contains 539 residues and 208 water molecules. The electron density is well-ordered except for region between Ala[40] and Ala[52] of subunit 3 that was modeled at 50% occupancy with a break at Lys[47]. This section of subunit 3 appears to adopt multiple conformations. In addition the electron density for the side chains
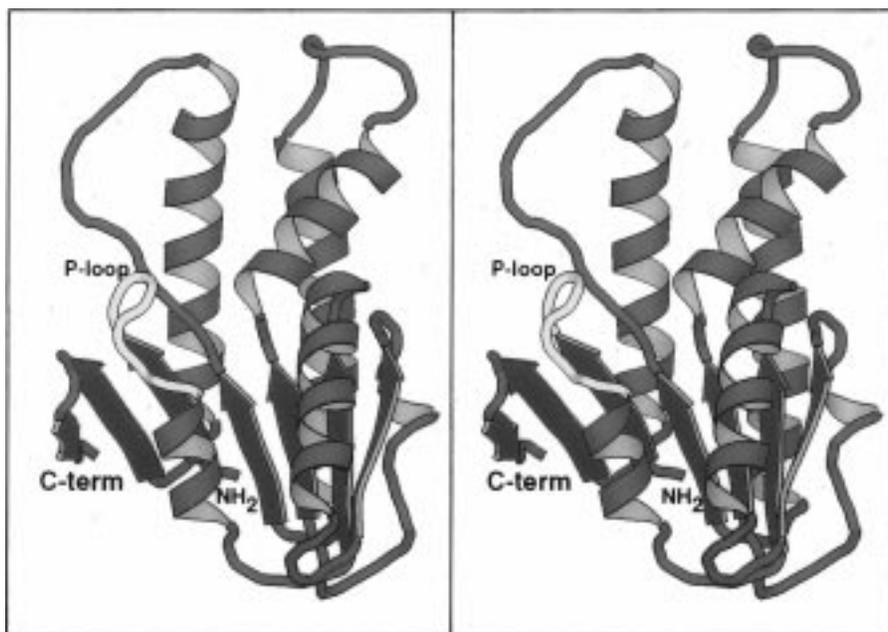
FIGURE 3: Stereo ribbon representation of one subunit of CobU where the P-loop is depicted in yellow. This figure and Figures 4, 6, and 7 were prepared with the program MOLSCRIPT (*58*).

of residues Asn[44] in subunit 1 and Arg[50] in all three subunits was disordered beyond the β-carbon; these residues and were modeled as alanines.

The rms differences in the positions of the α-carbons when the three independent subunits were superimposed were 0.41, 0.52, and 0.46 Å between subunits 1 and 2, 1 and 3, and 2 and 3, respectively. The respective rotational angles are 121.3°, 118.5°, and 120.2°.

During the refinement it was observed that chirality of Cys[81] in all subunits attempted to change. Examination of the electron density, conformational angles, and hydrogen-bonding patterns surrounding this amino acid residue suggested that the peptide bond between residues Glu[80] and Cys[81] adopts the cis configuration (Figure 2b,c). Once modeled in this state the refinement proceeded readily. Non-prolyl cis peptide bonds are rare in proteins (*30, 31*) and usually function to orient residues for ligand binding or catalysis (*32*). The cis peptide bond in CobU is found at the connection between a β-strand and an α-helix and is located opposite to the P-loop. This might provide an explanation for this configuration since it serves to orient the side chains of both Glu[80] and Cys[81] in the kinase active site. Indeed, the carboxylate group of Glu[80] is hydrogen-bonded to Oγ of Ser[13] of the P-loop. Both Ser[13] and Glu[80] are in a position where they might be expected to be involved in coordinating the divalent cation associated with the nucleotide; however, this conjecture can only be proven by determining the structure of CobU in the presence of nucleotides.

In addition to the protein and water molecules some form of polyvalent ion (most likely a phosphate or sulfate ion) is coordinated at half occupancy by the P-loop of each subunit. Although neither sulfate nor phosphate was included in any of the crystallization solutions, these ions are common contaminants of poly(ethylene glycol) (*33*). These ions are close to the position typically occupied by the β-phosphate of the triphosphate moiety in enzymes that utilize a P-loop to coordinate the nucleotide (*34*).

Tertiary Structure. The molecular architecture of the CobU subunit consists of a single domain with overall dimensions of approximately 41 Å × 48 Å × 48 Å (Figure 3). As can be seen, the structure is dominated by a seven-stranded mixed β-sheet flanked on either side by a total of five α-helices and one helical turn. These helices range in length from approximately 9 amino acid residues for the first α-helix at the N-terminus to about 20 amino acid residues for the last α-helix. Six of the seven β-strands run parallel. The C-terminal β-strand lies at the edge of the sheet and runs antiparallel to the others. The structure and topology of CobU is different from the cobalamin binding domains in either in methylmalonyl-CoA mutase (*18, 19*) and methionine synthase (*20*). As expected from amino acid sequence analyses, the CobU subunit contains a P-loop, or phosphate binding loop, defined by Met[1] to Gly[21]. This type of structural motif has been observed in many of nucleotide-dependent enzymes such as ATP synthase (*35, 36*), the G-proteins (*37, 38*), and myosin subfragment 1 (*39, 34*) and serves to properly position the triphosphate moiety of the nucleotide. Interestingly, the topology of CobU is identical to the central domain of the RecA protein which also contains a P-loop and binds ATP (*40, 41*).

The quaternary structure of CobU is trimeric, rather than dimeric as originally reported (*10*), where the molecule appears somewhat propellerlike with overall molecular dimensions of approximately 64 Å × 77 Å × 131 Å (Figure 4). This aggregation state is supported by equilibrium centrifugation studies that also show that the molecule is trimeric in solution (M. G. Thomas and J. C. Escalante-Semerena, unpublished results). The location of the P-loop is indicated in Figure 4. Although little is known about the location of the binding site for adenosylcobinamide or adenosylcobinamide phosphate, the shape of the molecule suggests that the binding site lies in the deep groove formed by the intersection of two adjacent subunits.

The subunit−subunit contacts of the trimer are formed primarily by three sections of the polypeptide chain: residues
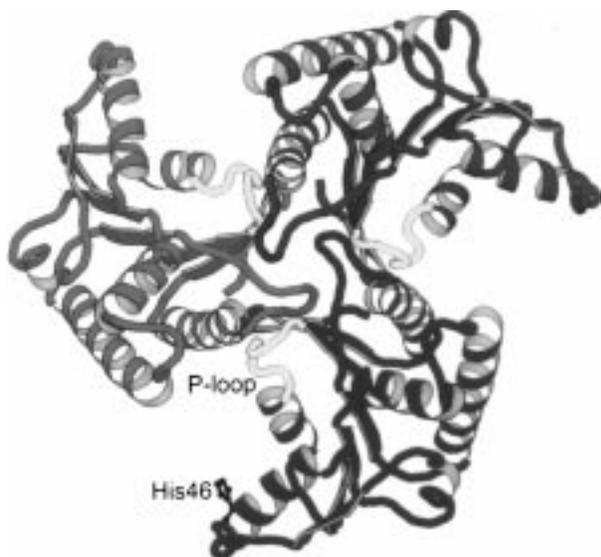
FIGURE 4: Ribbon representation of the quaternary structure of the CobU trimer viewed approximately parallel to the local 3-fold rotational axis. This reveals the location of the P-loop and His[46] that are the putative locations of the kinase and guanylyltransferase active sites.

Ala[8] to Gly[11], Glu[135] to approximately Ile[140], and Asp[152] to the C-terminus. Interestingly, Met[138] from each subunit abuts the 3-fold axis in a symmetric hydrophobic interaction. The hydrogen-bonding pattern between one subunit and its neighbors is very similar for all subunits, as would be anticipated from a 3-fold relationship between them. The total buried surface area of the trimer is ~4800 $Å^2$ or about 1600 $Å^2$/subunit. This represents ~17% of the surface area of the monomer, which is typical for a trimeric enzyme of this magnitude (*42*).

Nucleotide Binding Sites. As noted previously, CobU has two distinct enzymatic activities: kinase and guanylyltransferase. On account of the great preference for GTP in the transferase reaction and limited nucleotide specificity for the kinase reaction, it is presumed that these nucleotides bind at different places on the enzyme. It appears likely that the kinase activity is associated with the P-loop since this motif is commonly used for phosphoryl transfer reactions and because the guanylyltranferase activity proceeds via a covalent intermediate. The later would be inconsistent with the function of the P-loop.

It has been observed that the kinase activity for the enzyme from *S. typhimurium* can utilize a variety of nucleotides including ATP and GTP (*10*). In the equivalent enzyme from *P. denitrificans* (CobP), competition experiments have shown that the preferred nucleotide is ATP rather than GTP even though the kinase activity was activated by the presence of GTP (*43*). The situation for CobU appears to be similar since guanylylation of this enzyme also stimulates the kinase activity (*10*); however, preference for the nucleotide was not determined. Together, the preliminary kinetic and structural evidence suggests that the major determinant for the kinase activity lies in the triphosphate segment of the nucleotide rather than the base. At present it is unknown how the nucleotide binds to the P-loop; however, in comparison to other nucleotide binding proteins such as myosin, adenylate kinase, or RecA protein (described below), the nucleotide binding pocket is very open, which suggests that few interactions will be formed by the nucleoside and the protein in its present conformation. These observations are also consistent with the demonstration that pyrophosphate is an inhibitor of the enzyme activity (M. G. Thomas and J. C. Escalante-Semerena, unpublished results).

As noted under Experimental Procedures, a rhenium−ATP compound was utilized to prepare a heavy derivative in this structure determination. The final coordinates for CobU were used as a phasing model for the derivative data set. After refinement, this showed that the rhenium binding site was located approximately 16 Å from the P-loop; however, no electron density could be observed for the ATP moiety. Thus is unclear if the compound bound specifically to the nucleotide binding site or at some other adventitious location. Although this proved to be a successful derivative, it provides no information concerning the nucleotide binding site.

The guanylyltransferase activity proceeds via a covalent enzyme-guanylyl intermediate that forms in the absence of a cobinamide substrate (*10*). By analogy with most other nucleotide transferases, it is anticipated that the guanylylation of the enzyme would occur at a histidine side chain (*44*, *45*), although the guanylyltransfer in the mRNA capping enzyme proceeds via a guanylyllysine intermediate (*15*, *46*). Amino acid sequence comparisons of CobU with analogous proteins from *E. coli, P. denitrificans*, *Rhodobacter capsulatus*, and *Synechocystis* sp. indicate that there is only one conserved histidine in the sequence, namely, His[46]. Conversely there are no conserved lysines except for that located in the P-loop, which would be unlikely to participate in a guanylylation reaction on account of its highly polar environment (*45*), which supports the proposed argument that His[46] is the site of guanylylation. The conserved histidine lies at the end of an α-helix and is solvent-exposed (Figure 4) and is not associated with any obvious structural features that might provide specificity for the guanylyl moiety. This is in great contrast to the uridylyl binding site in galactose-1-phosphate uridylyltransferase, where the orientation and position of the histidine that forms the phosphoramidate intermediate is highly constrained by the protein tertiary structure (*45*). In addition, there is an obvious specificity pocket for the pyrimidine in galactose-1 phosphate uridylyltransferase. This suggests that CobU undergoes a conformational change prior to or during formation of the guanylyl intermediate in order to provide specificity for GTP.

Interestingly, the P-loop and the conserved histidine lie 21 Å apart and are located on opposite sides of the large cleft formed by the juxtaposition of adjacent subunits in the trimer (Figure 4). On account of the small size of the protein it seems reasonable that both enzyme activities utilize the same binding site for the corrin ring. If this is true, then the aminopropanol linker must swing or rotate from the kinase active site to the guanylylation site. Even if the aminopropanol linker does rotate in the active site it would appear that the protein must also undergo a conformational change when cobinamide binds or after guanylylation in order to bring the two active sites into closer proximity since in the native structure they are separated by more than twice the length of the linker. Conformational flexibility that serves to keep an enzymatic intermediate associated with a multi-enzyme complex has been suggested to occur in large complex protein systems, for example, in both acetyl-coenzyme A carboxylase and pyruvate dehydrogenase, for

```
        β₁              α₁            β₂              α₂
       ┌───┐        ┌──────────┐    ┌───┐        ┌──────────────┐
       2    6       12        20    25   30       33
S.t  ~~~~~~~MILVTGGARSGKSRHAEALIGDAPQVL.YIATS.QILDDEMAAR
E.c  ~~~~~~~MMILVTGGARSGKSRHAEALIGDSSQVL.YIATS.QILDDEMAAR
P.d  MSSLSAGPVLVLGGARSGKSSFSERLVEASGFTMHYVATG.RAWDDEMRER
R.c  ~~~~MAFTLLVTGGARSGKSSYAEARTLALGQPAHYIATS.EIWDDEMAER
S.   ~~~~MAFPILVTGPSRSGKSEWAEHLAAQNHPSVIYIATGKENAADVDQNR

            β₃          α₃              β₄          α₄
    ┌────┐  ┌──────┐  ┌────┐          ┌────┐  ┌──────────────┐
    47      53    58  63 65            76   80   83
S.t  IQHHKDGRPAHWRTAECWRHLDTLITADLAPDDAILLECITTMVTNLLFAL
E.c  IEHHRQGRPEHWRTVERWQHLDELIHADINPNEVVLLECVTTMVTNLLFDY
P.d  IDHHRTRRGEGWTTHEEPLDLVGILRRIDDPSHVVLIDCLTLWVTHLMLE.
R.c  IAAHQARRGPEWVTRAEPLDLVTALRETDDAPR..LIDCLTLWLTNLMLG.
S.   IQQHRDRRPSHWQTLCVDQNLTDTL.ETLPAGSCALVDSLGGWVANTL...

       ▽                   α₅              β₅
    ┌────┐          ┌────────────────────┐  ┌──────┐
    94                104            124  128 132
S.t  GGENDPEQWDYAAMERAIDDEIQILIAACQRCPAKVVLVTNEVGMGIVPEN
E.c  GGDKDPDEWDYQAMEQAINAEIQSLIAACQRCPAKVVLTNEVGMGIVPES
P.d  ..............ERDMTAEFAALVAYLPEARARLVFVSNEVGLGIVPEN
R.c  ..............DHDWEAAAKRLLAVLPELTAPVVFVTNEVGLGIVPDN
S.   ..ELSPEQWQTKQRE......LITFATQGGQNTPITIIFVAEETGWGVIPVY

             α₆                 β₆      β₇
    ┌──────────────────┐      ┌────┐  ┌──────┐
    145            164        167  172 175 178
S.t  RLARHFRDIAGRVNQRLAAAADEVWLVVSGIGVKIK*~~~~~~~~~~
E.c  RLARHFRDIAGRVNQQLAAAANEVWLVVSGIGVKIK~~~~~~~~~~~
P.d  RMAREFRDHAGRLHQIVAEKSAEVYFVAAGLPLKMKG~~~~~~~~~~
R.c  RLARAFRDNAGHLNQWVAAAADEVVFTVAGLPMKVK*~~~~~~~~~~
S.   ALGRIFRDRLGALTRQLGGLCGQVYLVTGGYAVDLRRWGEPLPPPRD
```

FIGURE 5: Sequence comparison of proteins showing significant sequence homology to CobU. The secondary structure for CobU is superimposed above the comparison. The sequences indicated as S.t, E.c, P.d, R.c, and S. correspond to the sequences from *Salmonella typhimurium*, *Escherichia coli*, *P. denitrificans*, *Rhodobacter capsulatus*, and *Synechocystis* sp., respectively. The completely conserved amino acids residues are highlighted in gray. The alignment was calculated with the GCG program package, version 8 (*60*).

which some structural information is available (*47, 48*). An alternative explanation for the distance between the P-loop and the conserved histidine might be that cobinamide and cobinamide phosphate might bind to the enzyme in different orientations in order to place the aminopropanol linker in proximity to the respective nucleotide. This latter possibility seems somewhat unlikely on account of the small size of the enzyme.

Relationship between CobU and Homologous Enzymes. Examination of the GenBank database with the sequence of CobU reveals four other sequences that exhibit significant similarity as detected with the program BLASTP (*49*). These are CobU from *Escherichia coli* (V33333), CobP from *P. denitrificans* (P29931), VbpT protein from *Rhodobacter capsulatus* (246611), and an undesignated protein from *Synechocystis* sp. (D64000) (*50–53*). The proteins from *E. coli*, *P. denitrificans*, *R. capsulatus,* and *Synechocystis* sp. share 82.2%, 42.7%, 52.2%, and 38.6% sequence identity, respectively, as shown by the sequence alignment in Figure 5. The locations of these identical residues mapped onto the structure of CobU (Figure 6) reveals that many of them are located surrounding the P-loop or are located on one face of the deep groove formed by the interface of the subunits. Since many of these conserved residues are solvent-exposed, this suggests that they serve a functional role and are most likely associated with substrate recognition. Of the conserved buried residues, the motif Gly[137]-X-Gly[139], where X is a methionine in CobU and a hydrophobic residue in all others, is of interest. This motif is located adjacent to the 3-fold axis where the methionines form a stacking interaction across the axis as noted earlier. Glycine residues are required in this location to allow the three chains to approach each other close to the axis and facilitate the interaction between the hydrophobic residues. The conformational angles (ϕ and ψ) for Gly[137] and Gly[139] are $-106°$ and $22°$ and $-114°$ and $-164°$, respectively, which are best suited to glycine residues.

Relationship between CobU and RecA Protein. Surprisingly, the RecA protein is the only protein at the present time in the Brookhaven Protein Data Bank, Accession Numbers 2REB and 1REA (*54, 41*), that shares a structural and topological similarity to CobU as determined with the program DEJAVU (*55*). The structural similarity between these proteins is shown in Figure 7 whereas the topological relationship is shown in Figure 8. As can be seen, CobU is remarkably similar to the central section of the RecA protein starting after the P-loop in both structures. The rms difference between 57 structurally equivalent α-carbon atoms in both structures is 1.23 Å as derived with the algorithm of Argos and Rossmann (*56*). There is no significant sequence similarity between these proteins except for the residues associated with the phosphate binding loop.
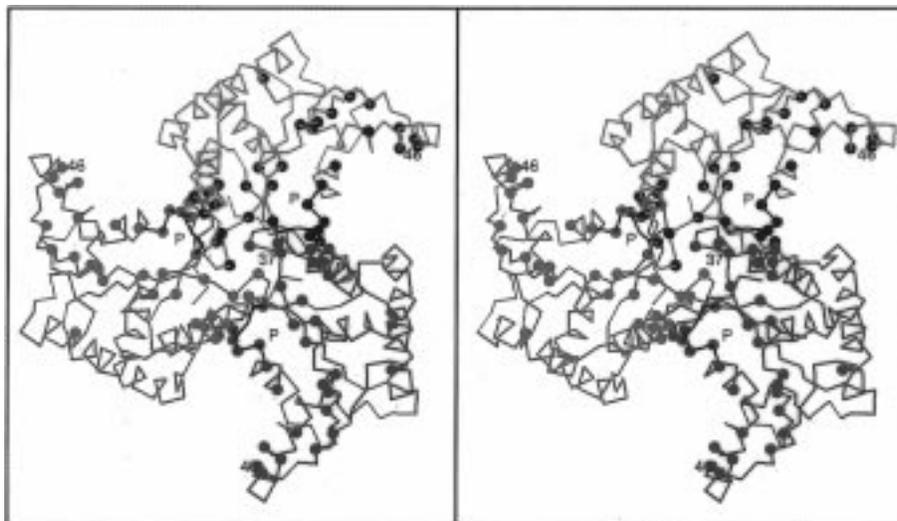

FIGURE 6: Stereoview of the location of the conserved residues mapped onto the tertiary structure of CobU. The P-loops are depicted in black.
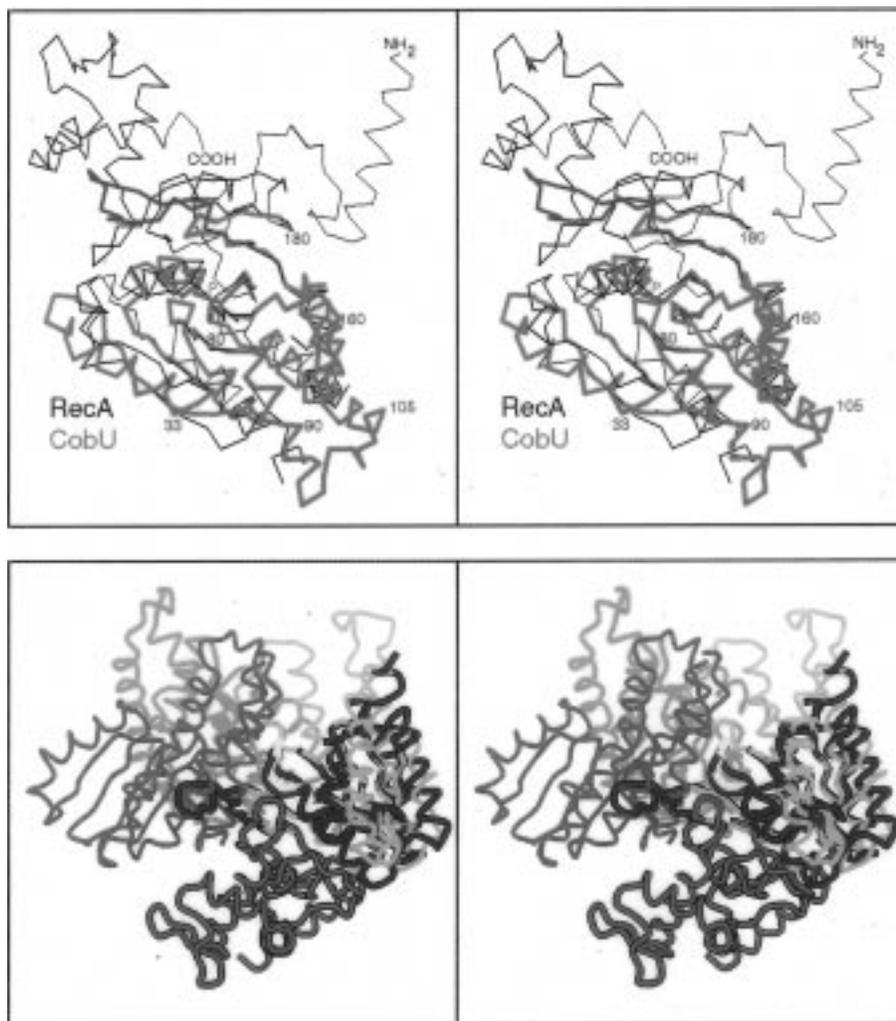
FIGURE 7: Structural comparison between CobU and RecA protein. (a, top) Stereo overlap of a CobU monomer and RecA protein. The longer loop in RecA protein between residues 228 and 240, which would interfere with trimer formation in CobU, is depicted in green. (b, bottom) Stereo overlap of RecA onto a single subunit of the CobU trimer where the P-loop of CobU is depicted in yellow. The proteins were aligned with the program OVRLAP (*56*).

RecA protein is considerably larger than CobU, where there are approximately 61 and 100 extra amino acid residues at the N- and C-termini, respectively, relative to CobU. In the central section of the RecA protein the lengths of the helices and connecting loops are quite similar to those found in CobU. There is one notable difference in the loop that connects the last two strands in CobU. In the RecA protein the structurally equivalent loop that joins the β-strands is fairly long and is formed by 13 amino acid residues (228–240). Conversely, in CobU the corresponding connection is made by a hairpin turn built from three amino acid residues. This difference arises because the loop in CobU abuts a 3-fold related subunit such that a longer loop could not be accommodated in this position without a substantial change in the structure of the protein (Figure 7b). Interestingly, the N- and C-terminal extensions of RecA protein relative to CobU are oriented away from the subunit–subunit interfaces in CobU (Figure 7b).

The similarity between CobU and RecA protein provides some insight into the kinase binding site. Superposition of ADP bound to RecA protein (*40*) onto CobU shows, as mentioned earlier, that the active site for CobU is very open, such that there are very few contacts between the protein and the substrate. The orientation of the diphosphate moiety

in RecA places restraints on the likely position of the γ-phosphate in CobU and hence on the direction of attack of the aminopropanol group that becomes phosphorylated. In turn this places restraints on the likely binding sites for the corrinoid. Together these considerations suggest that a conformational change must occur when nucleotide and/or cobinamide binds to the enzyme.

The meaning of the similarity between CobU and RecA protein is unclear. There is no obvious relationship between the reactions that these proteins catalyze except that they both utilize ATP. Even here there are differences, since with RecA hydrolysis of ATP is involved in driving DNA strand exchange (*57*), whereas CobU functions as a kinase. The obvious structural similarity between these proteins suggests that they might have evolved from a common ancestor, although it is hard to suggest what the initial function of that protein might have been. Conversely, these proteins might represent an example of the same fold originating within two primordial proteins of different function.

CONCLUSIONS

The structure of adenosylcobinamide kinase/adenosylco- binamide phosphate guanylyltransferase presented here an-
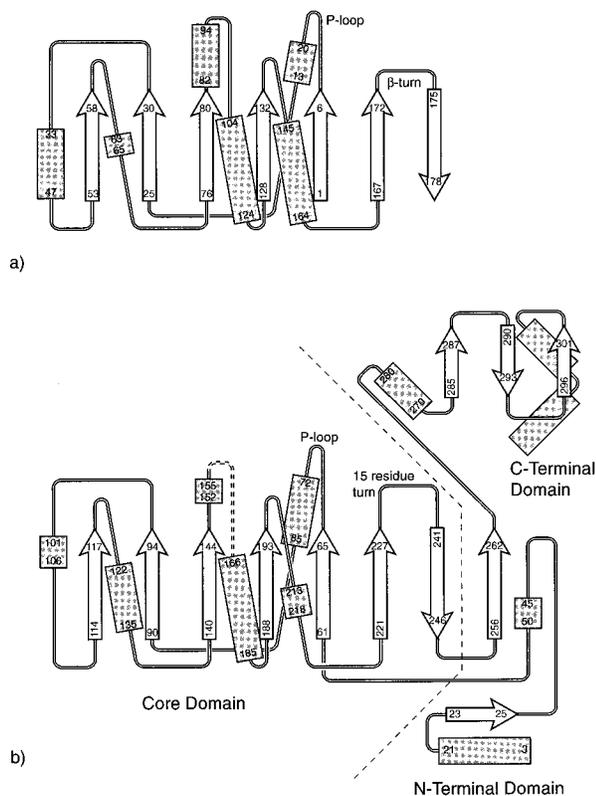
FIGURE 8: Topology drawing for (a) CobU and (b) the RecA protein.

swers some of the questions concerning the architecture of the enzyme and disposition of its putative enzyme activities. Interestingly, CobU does not share any structural similarity with the cobalamin binding domains of either methylmalonyl-CoA mutase (*18*, *19*) or methionine synthase (*20*), which was expected since the chemistry of CobU is associated with the aminopropanol side chain rather than the central cobalt atom of the corrin ring as with the latter two enzymes. The structure of CobU reveals that the molecule is a molecular trimer where the putative kinase and transferase active sites are located in a deep groove formed by the association of two neighboring subunits. Although the position of the binding site for the corrin ring is unknown, it seems likely that the binding site is associated with this deep groove. This conjecture is supported by the numerous highly conserved amino acid residues that line the exposed surfaces of this structural feature. If it is assumed that the enzyme has only one binding site for adenosylcobinamide, the observed 21 Å separation between the putative kinase and guanylyltransferase sites requires that the protein undergo a conformational change during catalysis in order to bring these sites closer together. This would be necessary since the phosphate group added to the aminopropyl arm of the cobinamide by the kinase activity carries out a nucleophilic attack on the enzyme-guanylyl intermediate to generate adenosylcobinamide-GDP. The aminopropyl arm is too short to simply rotate from one site to the other without a conformational change. Earlier kinetic studies of CobU and its homologue in *P. denitrificans* demonstrated that guanylylation of the enzyme activates the kinase site (*43*, *10*), which cannot be explained by the present structure unless a conformational change is invoked during the formation of the nucleotidyl intermediate. Thus far it has been difficult to study the

kinase and transferase activities independently because of their influence on each other. The current structure provides guidance on how to distinguish between the two enzymatic activities by site-directed mutagenesis. The structure also indicates that the nature of the cobinamide binding site and the communication between the two active sites can only be understood by further structural analyses. These investigations are in progress.

REFERENCES

1. Lenhert, P. G., and Hodgkin, D. C. (1961) *Nature 192*, 937−938.
2. Scott, A. I. (1993) *Angew. Chem., Int. Ed. Engl. 32*, 1223−1243.
3. 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.
4. O'Toole, G. A., Rondon, M. R., Suh, S.-J., Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1996) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F. C., Ed.), American Society for Microbiology Press, Washington, D.C.
5. Rondon, M. R., Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1997) *Prog. Nucleic Acid Res. Mol. Biol. 56*, 347−384.
6. Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1997) *J. Biol. Chem. 272*, 17662−17667.
7. O'Toole, G. A., Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1994) *J. Biol. Chem. 269*, 26503−26511.
8. O'Toole, G. A., Rondon, M. R., and Escalante-Semerena, J. C. (1993) *J. Bacteriol. 175*, 3317−3326.
9. Roth, J. R., Lawrence, J. G., Rubenfield, M., Kieffer-Higgins, S., and Church, G. M. (1993) *J. Bacteriol. 175*, 3303−3316.
10. O'Toole, G. A., and Escalante-Semerena, J. C. (1995) *J. Biol. Chem. 270*, 23560−23569.
11. O'Toole, G. A. (1994) Ph.D. Thesis, University of Wisconsin, Madison, WI.
12. Brushaber, K., O'Toole, G. A., and Escalante-Semerena, J. C. (1998) *J. Biol. Chem. 273*, 2684−2691.
13. Martin, S. A., and Moss, B. (1975) *J. Biol. Chem. 250*, 9330−9335.
14. Shuman, S., and Hurwitz, J. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 187−191.
15. Shuman, S., and Schwer, B. (1995) *Mol. Microbiol. 17*, 405−410.
16. Szumilo, T., Drake, R. R., York, J. L., and Elbein, A. D. (1993) *J. Biol. Chem. 268*, 17943−17950.
17. Liu, J. J., and McLennan, A. G. (1994) *J. Biol. Chem. 269*, 11787−11794.
18. Cannata, J. J. B., Focesi, A. J., Mazumder, R., Warner, R., and Ochoa, S. (1965) *J. Biol. Chem. 240*, 3249−3257.
19. 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.
20. Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G., and Ludwig, M. L. (1994) *Science 266*, 1669−1674.
21. Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 67−71.
22. Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 916−924.
23. Wesenberg, G., and Rayment, I. (1997) manuscript in preparation.

24. Terwilliger, T. C., and Eisenberg, D. (1983) *Acta Crystallogr. A39*, 813−817.
25. Rypniewski, W. R., Breiter, D. R., Benning, M. M., Wesenberg, G., Oh, B. H., Markley, J. L., Rayment, I., and Holden, H. M. (1991) *Biochemistry 30*, 4126−4131.
26. Bricogne, G. (1976) *Acta Crystallogr. A32*, 832−847.
27. Jones, T. A. (1985) in *Methods in Enzymology* (Wycoff, H. W., Hirs, C. H. W., & Timasheff, S. N., Eds.) Vol. 115, pp 157−171, Academic Press Inc., New York.
28. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489−501.
29. Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
30. Ramachandran, G. N., and Mitra, A. K. (1976) *J. Mol. Biol. 107*, 85−92.
31. Stewart, D. E., Sarkar, A., and Wampler, J. E. (1990) *J. Mol. Biol. 214*, 253−260.
32. Herzberg, O., and Moult, J. (1991) *Proteins: Struct., Funct., Genet. 11*, 223−229.
33. Jurnak, F. (1985) *J. Mol. Biol. 185*, 215−217.
34. Smith, C. A., and Rayment, I. (1996) *Biophys. J. 70*, 1590−1602.
35. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) *EMBO J. 1*, 945−951.
36. Abrahams, J. P., Leslie, A. G. W., Lutter, R., and Walker, J. E. (1994) *Nature 370*, 621−628.
37. Pai, E. F., Kabsch, W., Krengel, U., Holmes, K. C., John, J., and Wittinghofer, A. (1989) *Nature 341*, 209−214.
38. Milburn, M. V., Tong, L., DeVos, A. M., Brünger, A., Yamaizumi, Z., Nishimura, S., and Kim, S.-H. (1990) *Science 247*, 939−945.
39. Rayment, I., Rypniewski, W. R., Schmidt-Bäse, K., Smith, R., Tomchick, D. R., Benning, M. M., Winkelmann, D. A., Wesenberg, G., and Holden, H. M. (1993) *Science 261*, 50−58.
40. Story, R. M., and Steitz, T. A. (1992) *Nature 355*, 374−376.
41. Story, R. M., Weber, I. T., and Steitz, T. A. (1992) *Nature 355*, 318−325.
42. Miller, S., Lesk, A. M., Janin, J., and Chothia, C. (1987) *Nature 328*, 834−836.
43. Blanche, F., Debussche, L., Famechon, A., Thibaut, D., Cameron, B., and Crouzet, J. (1991) *J. Bacteriol. 173*, 6052−6057.
44. Wong, L.-J., Sheu, K.-F. R., Lee, S.-L., and Frey, P. A. (1977) *Biochemistry 16*, 1010−1016.
45. Wedekind, J. E., Frey, P. A., and Rayment, I. (1996) *Biochemistry 36*, 11560−11569.
46. Håkansson, K., Doherty, A. J., Shuman, S., and Wigley, D. B. (1997) *Cell 89*, 545−553.
47. Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A., and Hol, W. G. (1992) *Science 255*, 1544−1550.
48. Waldrop, G. L., Rayment, I., and Holden, H. M. (1994) *Biochemistry 33*, 10249−10256.
49. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403−410.
50. Crouzet, J., Levy-Schil, S., Cameron, B., Cauchois, L., Rigault, S., Rouyez, M. C., Blanche, F., Debussche, L., and Thibaut, D. (1991) *J. Bacteriol. 173*, 6074−6087.
51. Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Sazuka, T., Miyajima, N., Sugiura, M., and Tabata, S. (1995) *DNA Res. 2*, 153−166.
52. Lawrence, J. G., and Roth, J. R. (1995) *J. Bacteriol. 177*, 6371−6380.
53. Pollich, M., and Klug, G. (1995) *J. Bacteriol. 177*, 4481−4487.
54. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rogers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.
55. Kleywegt, G. J., and Jones, A. T. (1997) in *Methods in Enzymology* (Carter, C. W. J., Sweet, R. M., Abelson, J. N., & Simon, M. I., Eds.) Vol. 277, pp 525−545, Academic Press, New York.
56. Rossmann, M. G., and Argos, P. (1975) *J. Biol. Chem. 250*, 7525−7532.
57. MacFarland, K. J., Shan, Q., Inman, R. B., and Cox, M. M. (1997) *J. Biol. Chem. 272*, 17675−17685.
58. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
59. Fisher, A. J. (1996) University of Wisconsin, Madison, WI.
60. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acid Res. 12*, 387−395.

BI973178F