# Common Chelatase Design in the Branched Tetrapyrrole Pathways of Heme and Anaerobic Cobalamin Synthesis†

Heidi L. Schubert,‡ Evelyne Raux,§ Keith S. Wilson,*,‡ and Martin J. Warren§

*Department of Chemistry, University of York, York YO10 5DD, U.K., and Department of Molecular Genetics, Institute of Ophthalmology, University College London, 11-43 Bath Street, London EC1V 9EL, U.K.*

*Received March 23, 1999; Revised Manuscript Received May 17, 1999*

ABSTRACT: Prosthetic groups such as heme, chlorophyll, and cobalamin (vitamin $B_{12}$) are characterized by their branched biosynthetic pathway and unique metal insertion steps. The metal ion chelatases can be broadly classed either as single-subunit ATP-independent enzymes, such as the anaerobic cobalt chelatase and the protoporphyrin IX (PPIX) ferrochelatase, or as heterotrimeric, ATP-dependent enzymes, such as the Mg chelatase involved in chlorophyll biosynthesis. The X-ray structure of the anaerobic cobalt chelatase from *Salmonella typhimurium*, CbiK, has been solved to 2.4 Å resolution. Despite a lack of significant amino acid sequence similarity, the protein structure is homologous to that of *Bacillus subtilis* PPIX ferrochelatase. Both enzymes contain a histidine residue previously identified as the metal ion ligand, but CbiK contains a second histidine in place of the glutamic acid residue identified as a general base in PPIX ferrochelatase. Site-directed mutagenesis has confirmed a role for this histidine and a nearby glutamic acid in cobalt binding, modulating metal ion specificity as well as catalytic efficiency. Contrary to the predicted protoporphyrin binding site in PPIX ferrochelatase, the precorrin-2 binding site in CbiK is clearly defined within a large horizontal cleft between the N- and C-terminal domains. The structural similarity has implications for the understanding of the evolution of this branched biosynthetic pathway.

Cofactors and prosthetic groups such as heme, chlorophyll, cobalamin (vitamin $B_{12}$), siroheme, and coenzyme $F_{430}$ are characterized by their tetrapyrrole-derived framework and a centrally chelated metal ion (*1*). The common macrocyclic structure is a consequence of an initial shared biosynthetic pathway, which branches to generate the various modified tetrapyrroles. The central metal ion is inserted by a chelatase that is specific to each individual branch (Figure 1).

Separate ferro, magnesium, and cobalt chelatases have been identified as a consequence of functional and genetic studies, as well as through genome sequencing projects. On the basis of limited biochemical characterization and sequence similarity, the chelatases appear to fall into two classes. The first class is composed of ATP-dependent heterotrimeric complexes, exemplified by both the chlorophyll/bacteriochlorophyll magnesium inserting enzyme and the aerobic cobalamin biosynthetic cobalt chelatase. For magnesium chelatase, this complex is made up of BchlD, -H, and -I (*2*), while the aerobic cobalt chelatase is comprised of CobN, -S, and -T subunits (*3*). There is significant sequence similarity between BchlH and CobN, both of which are responsible for substrate binding, but no similarity exists between any of the remaining subunits except for a putative ATP binding site located on BchlI and CobS.

Members of the second class of chelatase are ATP-independent, single-subunit, enzymes as exemplified by protoporphyrin IX (PPIX)[1] ferrochelatase, an enzyme that inserts ferrous ions to yield protoheme (*4*). Other members of this class include the sirohydrochlorin ferrochelatase of siroheme synthases, CysG (*5*) and Met8p (*6*), and the anaerobic cobalt chelatases, CbiK (*7*) and CbiX (*8, 9*). There is no significant sequence similarity between any of these chelatases.

The mechanism of the anaerobic cobalt chelatase has not been extensively studied, but is expected to be similar to that suggested for PPIX ferrochelatase (*10*−*12*). A distortion of the bound porphyrin is required to expose the nitrogen lone-pair electrons to the incoming metal ion. A general base is required to abstract protons from the ring nitrogens prior to iron insertion, postulated to be a glutamic acid residue in ferrochelatase (*13*) (Glu264; all PPIX ferrochelatase numbering will refer to the *Bacillus subtilis* enzyme). The major protein ligand for the ferrous ion is a histidine residue (His183) (*14*). Evidence for this mechanism has come from spectroscopic investigations, site-directed mutagenesis studies, and the recent structure determination of *B. subtilis* PPIX ferrochelatase (*15*). Glu264 and His183 are mechanistically the most critical residues for catalytic activity and form the active site in conjunction with residues that selectively bind the tetrapyrrole substrate. It has been suggested that proto-

† We gratefully acknowledge funding from the National Institutes of Health Diabetes, Digestive and Kidney Disease Institute (DK09549-01X1), the Wellcome Trust (047293), and Biotechnology and Biological Sciences Research Council Grants SB09829 and BO8486/BO8652.

* Correspondence should be addressed to this author. Phone: 44-1904-432519. Fax: 44-1904-410519. E-mail: keith@yorvic.york.ac.uk.

‡ University of York.

§ University College London.

[1] Abbreviations: PPIX, protoporphyrin IX; IPTG, isopropyl-$\beta$-D-thiogalactopyranoside; rms, root-mean-square; mm, minimal media; MS, methionine synthase; MMCoAM, methylmalonyl-CoA mutase; *S. typhimurium*, *Salmonella typhimurium*; *B. subtilis*, *Bacillus subtilis*; *P. denitrificans*, *Para denitrificans*.

10.1021/bi9906773 CCC: $18.00 
Published on Web 07/28/1999

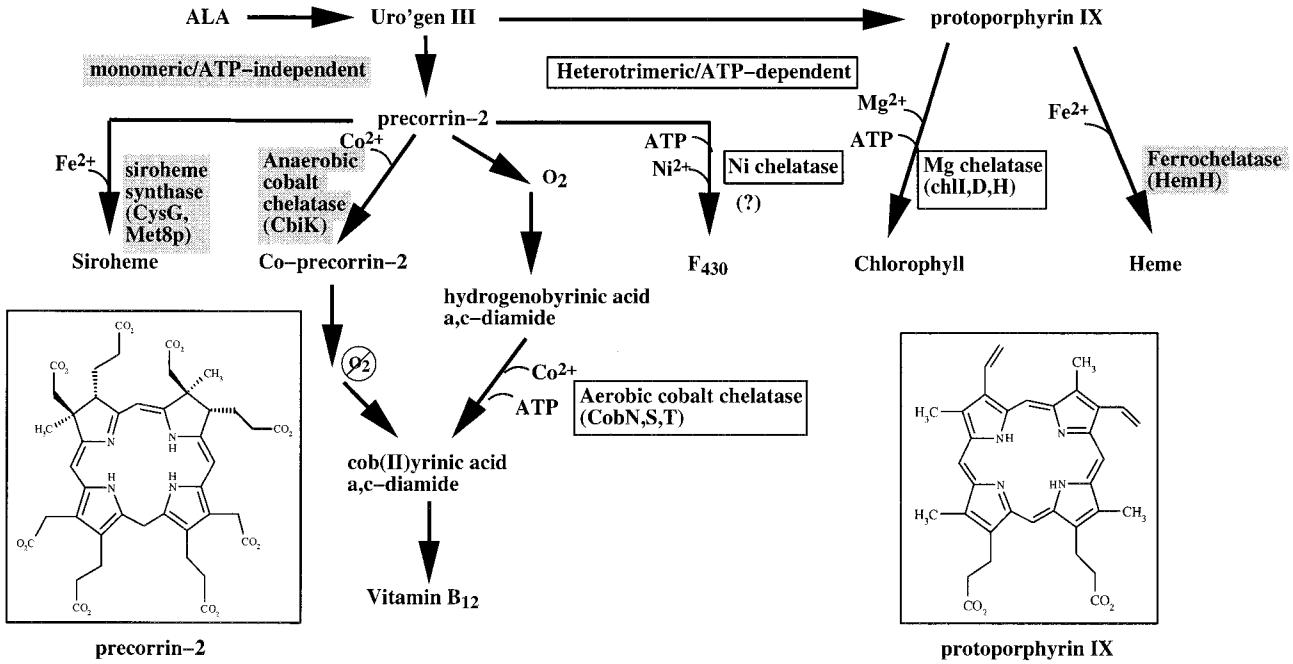


FIGURE 1: Tetrapyrrole biosynthesis. Diagramatic representation of the relationship between the tetrapyrrole biosynthetic pathways, highlighting the common metal ion chelation step. The chelatases can be separated into two classes of enzymes: (1) single-subunit ATP-independent enzymes such as the anaerobic cobalt chelatase and the PPIX ferrochelatase; and (2) the heterotrimeric ATP-dependent chelatases as exemplified by the aerobic cobalt chelatase and protoporphyrin monomethyl ester magnesium chelatase.

porphyrin binds between Glu264 and His183 although this is not feasible without a significant conformational change in the enzyme (*15*).

In this paper, we describe the structure of the anaerobic cobalamin biosynthetic cobalt chelatase, CbiK, from *S. typhimurium* and show that it has a remarkably similar topology to *B. subtilis* ferrochelatase despite minimal amino acid conservation. Based on this structure, several key residues within CbiK have been mutated to test their involvement in the catalytic mechanism. The importance of the structural similarity is discussed with respect to the mechanism and the evolution of the two proteins.

## MATERIALS AND METHODS

*Purification, Crystallization, and Structure Determination.* Recombinant histidine-tagged *S. typhimurium* CbiK was overexpressed in *E. coli* BL21(DE3)pLysS cells by cloning the gene into pET14b (*9*) and purified using nickel-charged chelating Sepharose (Pharmacia) (*16*). The purified protein was dialyzed against 20 mM sodium citrate, pH 6.5, and 100 mM NaCl. The buffer was adjusted for thrombin cleavage of the histidine-tag by the addition of 70 mM Tris, pH 8.5, and 2.5 mM $CaCl_2$. The protein was incubated with thrombin (10 $\mu$g/mL) overnight at room temperature. The cleaved protein was purified by gel filtration on a Sephadex G75 16/60 column equilibrated in the dialysis buffer, concentrated to 7 mg/mL, and stored at 4 °C for use in crystallization trials.

Hexagonal crystals were grown using the hanging-drop method from both His-tagged and cleaved protein concentrated to 7 mg/mL and mixed in equal volume with 10−15% poly(ethylene glycol) (MW 4000), 0.1 M Tris buffer, pH 8.5, and 0.2 M $Li_2SO_4$ and equilibrated over the same solution. This solution, supplemented with 30% glycerol, served as a cryoprotectant for X-ray data collection at 120 K on an R-axis area detector using copper radiation, $\lambda$ = 1.54 Å, from an in-house rotating-anode X-ray source. The crystals are in space group $P6_322$ with cell dimensions $a = b = 128.08$ Å, $c = 85.44$ Å and diffract to 2.4 Å. The overall temperature factor (*17*) of the data is quite high at 36.3 Å$^2$, and the mosaicity of the crystals is ~0.4°. Data were indexed and scaled using DENZO and SCALEPACK (*18*). There is one molecule in the asymmetric unit, and the Matthews coefficient ($V_m$) of 3.4 Å$^3$/Da indicates a solvent content of ~63% (*19*). Heavy-atom derivatives were obtained by soaking the crystals for 1 h prior to data collection in solutions equivalent to the mother liquor supplemented with 1 mM of the heavy atom compounds, MeHgCl or $(Me)_3$-PbOAc.

The structure of CbiK was determined by multiple isomorphous replacement with anomalous scattering. Two single-site heavy-atom derivatives were identified by difference Patterson methods using the CCP4 package (*20*). The location of the heavy atom is approximately the same for each derivative and was refined in MLPHARE (Table 1). The correlation coefficient of the MIRAS phases before and after solvent flattening (60% solvent) was 0.48/0.76 (*21*). The entire structure was built into the solvent-flattened map using QUANTA (*22*) and refined using REFMAC (*23*) against all data between 20 and 2.4 Å resolution minus a 'random' set of 5% of the data for $R_{free}$ calculation. The final model contains 257 residues of CbiK (2−258), 3 sulfate ions, and 224 waters. No density is available for the first residue of CbiK, the residual N-terminal tag, or the last six C-terminal residues of the protein. The *R*-factor for this model is 19.9%, and the $R_{free}$ is 26.9% to 2.4 Å resolution with good stereochemistry (Table 2). Coordinate and diffraction data have been deposited in the Brookhaven Protein Database (PDB code: 1QGO) and are available from the authors.

Table 1: Data Collection Statistics

| | native CbiK[a] | MeMgCl[a] | $(Me)_3PbOAc$[a] | *F222* |
|---|---|---|---|---|
| $a = b = c =$ (Å) | 128.08, 85.44 | 127.62, 84.90 | 127.50, 85.23 | 65.8, 123.2, 137.7 |
| observed reflections | 16699 | 13805 | 11694 | 59038 |
| unique reflections | 16597 | 13702 | 11515 | 3072 |
| resolution (Å) (last shell) | 30−2.4 (2.49−2.4) | 30−2.55 (2.64−2.55) | 30−2.7 (2.8−2.7) | 20−3.5 (3.6−3.5) |
| completness (%) (last shell) | 99.5 (99.8) | 99.4 (99.9) | 98.6 (99.5) | 83 (86) |
| $R_{merge}$[b] (last shell) | 8.7 (31) | 9.4 (26) | 11.9 (41) | 13.2 (34) |
| phasing power (acentric/centric) | − | 1.20/0.74 | 1.72/1.15 | − |

[a] Space group $P6_322$. [b] $R_{merge} = 100\sum |I - \langle I\rangle|/\sum I$, where $I$ is an individual intensity measurement and $\langle I\rangle$ is the mean intensity for symmetry-related reflections.

Table 2: Refinement Statistics

| | CbiK |
|---|---|
| resolution range (Å) | 20.0−2.4 |
| no. of protein atoms[a] | 2014 |
| no. of solvent atoms[a] | 224 |
| *R*-factor (%)[b] | 19.56 |
| $R_{free}$ (%)[c] | 26.34 |
| Δ bond lengths (Å)[d] | 0.012 |
| Δ bond angles (deg)[d] | 2.0 |
| $\langle B\rangle$ (Å$^2$) main chain[e] | 34.23 |
| $\langle B\rangle$ (Å$^2$) side chain[e] | 37.20 |
| $\langle B\rangle$ (Å$^2$) water molecules[e] | 47.86 |
| $\langle B\rangle$ (Å$^2$) $SO_4$[d] | 42.47 |
| no. of $\Phi/\psi$ angles (%) in most favored positions (allowed) | 93.9 (6.1) |

[a] Non-hydrogen atoms only. [b] *R*-factor $= 100\sum(||F_o| - |F_c||/|\sum F_o|$, where $F_o$ and $F_c$ are the observed and calculated structure factors. [c] $R_{free}$ = *R*-factor is calculated for a selected subset of the reflections (5%) excluded from the refinement. [d] rms deviation from ideal values. [e] Wilson *B* of the data is 36.3 (*17*).

A second crystal form was obtained in the same manner by incubating the cleaved CbiK with an equal volume of 250−300 mM sodium acetate, 100 mM sodium citrate buffer, pH 6.5, and 2 mM dithiothreitol. The resultant thin flat plates are in space group *F222* with cell dimensions $a = 65.8$ Å, $b = 123.2$ Å, $c = 137.7$ Å but only diffract to 3.5 Å resolution (Table 1). There is one molecule in the asymmetric unit and 46% solvent with a Matthews coefficient of 2.3 Å$^3$/Da. The structure of these crystals was solved by molecular replacement using the $P6_322$ model but is insufficiently defined for critical evaluation. Both crystal forms could be grown from protein containing the histidine tag, or with the tag cleaved off. In addition, both crystal forms can be grown in the same crystallization experiment, though the best conditions are as stated above. Attempts to obtain either a cobalt derivative or a precorrin-2-bound complex have thus far been unsuccessful.

*Site-Directed Mutagenesis.* Five individual CbiK mutations, H145A, H207A, D211A, E175A, and E89A, and a double mutant, H145A/H207A, were produced by site-directed mutagenesis using the Sculptor *in vitro* mutagenesis kit RPN1526 from Amersham. All mutations were verified through DNA sequencing. The *cbiK* gene and its mutated variants were subsequently cloned into pER119 (pKK223-3/*cobA*) 5′ of the *cobA* gene of *P. denitrificans* to yield pER303 and its derivatives (Table 3). When transformed into an *E. coli cysG* deletion strain (ER171), the pER303 series of plasmids were used to investigate the *in vivo* functionality of *cbiK* as described below.

*Biological Characterization of Mutants.* ER171 (*cysG* deletion strain) was individually transformed with pER119 and each of the pER303 family members. The resulting strains (ER316−323, Table 3) were streaked on minimal medium (mm) plates, mm-plates containing cobalt and mm-plates containing cysteine. Table 4 shows the results obtained after 40 h growth at 37 °C.

Cell extracts of the eight strains (ER316−323) described in the previous section were grown anaerobically in minimal medium supplemented with cysteine, aminolevulinic acid, and IPTG. All strains were grown in the presence or absence of exogenous cobalt. The accumulated modified tetrapyrroles were isolated on a DEAE-Sephacel column after cell lysis, and were eluted in 50 mM Tris, pH 7.5, containing 1 M NaCl. A UV/VIS spectrum was immediately recorded.

An *E. coli cysG* deleted strain, containing all the genetic information to synthesize cobyric acid from cobalt precorrin-2 except for *cbiK* (ER185), was separately transformed with pER119 and the pER303 family of plasmids. The resulting strains, ER306−313, were grown anaerobically in the presence or absence of exogenous cobalt, and the quantity of cobyric acid produced was estimated from a modified bioassay (*24*).

## RESULTS AND DISCUSSION

*CbiK Structure.* The anaerobic cobalt chelatase, CbiK, is a bilobal enzyme containing two $\alpha/\beta$ domains related to one another by a pseudo-2-fold symmetry, similar to the fold found in the periplasmic-type protein family (*25*). Each domain of CbiK contains a four-stranded parallel $\beta$-sheet (strand order: 2134) surrounded by at least four helices (Figure 2A,B). Helices α4 and α8 link the domains such that the N-terminal domain can be considered to contain helices α1−3 and α8 while the C-terminal domain contains helices α4−7. The N-terminal domain contains an additional helix inserted prior to α2, which is labeled α1′ to distinguish it as an additional helix that does not obey the pseudo-2-fold symmetry. In the C-terminal domain, four residues interrupt the secondary structure of α7, but the helix continues after a bulge. The topological similarity between the domains (4.3 Å rms deviation between domains over 105/135 residues (*26*)) suggests that they have arisen through an evolutionary gene duplication event.

Analytical gel-filtration studies show that CbiK elutes with a molecular mass of ∼80 kDa, suggesting that the native enzyme is a homotrimer of subunit molecular mass 30 kDa (*7*). A putative trimer is generated by the 3-fold axis of the $P6_322$ crystal form, but the alternate *F222* form does not contain a 3-fold axis and contains a monomeric form of the enzyme. Though the stated crystallographic conditions for each crystal form are different, frequently both forms appear in the same experiment, suggesting that there is an equilib-

Table 3: List of Strains and Plasmids

| plasmids | properties |
|---|---|
| pAR8086 | pACYC184/*lacI*$^q$ (ref *24*) |
| pER119 | pKK223.3 derived/*P. denitrificans cobA* (*7*) |
| pER126$^\Delta$ | pACYC184/*S. typhimurium-cbiA-C-D-E-T-F-G-H-J-K*$^\Delta$*-L-M-N-Q-O-P* (*7*) |
| pER231 | pET14b/*S. typhimurium cbiK* (*9*) |
| pER303 | pKK223.3 derived/*S. typhimurium cbiK* and *P. dentrificans cobA* |
| strains | properties |
| ER171 | 302Δa (pAR8086) (*7*) *E. coli cysG* deleted strains with pACYC184-*lacI*$^q$ |
| ER227 | BL21(DE3)(pLysS)(pER231) (*7*) overexpresses His-tagged *S. typhimurium cbiK* |
| ER306 | 302Δa (pER126K$^\Delta$)(pER119) |
| ER307 | 302Δa (pER126K$^\Delta$)(pER303) |
| ER308 | 302Δa (pER126K$^\Delta$)(pER303$^{E89A}$) |
| ER309 | 302Δa (pER126K$^\Delta$)(pER303$^{H145A}$) |
| ER310 | 302Δa (pER126K$^\Delta$)(pER303$^{E175A}$) |
| ER311 | 302Δa (pER126K$^\Delta$)(pER303$^{H207A}$) |
| ER312 | 302Δa (pER126K$^\Delta$)(pER303$^{D211A}$) |
| ER313 | 302Δa (pER126K$^\Delta$)(pER303$^{H145A/H207A}$) |
| ER316 | 302Δa (pAR8086)(pER119) |
| ER317 | 302Δa (pAR8086)(pER303) |
| ER318 | 302Δa (pAR8086)(pER303$^{E89A}$) |
| ER319 | 302Δa (pAR8086)(pER303$^{H145A}$) |
| ER320 | 302Δa (pAR8086)(pER303$^{E175A}$) |
| ER321 | 302Δa (pAR8086)(pER303$^{H207A}$) |
| ER322 | 302Δa (pAR8086)(pER303$^{D211A}$) |
| ER323 | 302Δa (pAR8086)(pER303$^{H145A/H207A}$) |

Table 4: CysG Complementation and Cobyric Acid Assays for Wild-Type CbiK and Mutant Constructs[a]

| plasmid pKK derived constructs | growth on minimal media after 40 h at 37 °C | | cobyric acid (pmol/$OD_{600}$) with $CoCl_2$, $6H_2O$ | |
|---|---|---|---|---|
| | no addition | $CoCl_2$ | 0 | 1 mg/L |
| *cobA* | − | − | 0 | 0 |
| *cobA-cbiK* | ++ | − | 9 | 100 |
| *cobA-cbiK*$^{E89A}$ | +++ | − | 7 | 136 |
| *cobA-cbiK*$^{H145A}$ | + | ± | 0 | 26 |
| *cobA-cbiK*$^{E175A}$ | +++ | − | 9 | 103 |
| *cobA-cbiK*$^{H207A}$ | +++ | ++ | 1 | 111 |
| *cobA-cbiK*$^{D211A}$ | ++ | − | 9 | 103 |
| *cobA-cbiK*$^{H145A/H207A}$ | − | − | 0 | 0 |

[a] All the strains for cobyric acid production are derived from strain ER185: *E. coli cysG* deleted strain (302Δa) containing the *S. typhimurium cbiA-C-D-E-T-F-G-H-J-K*$^\Delta$*-L-M-N-Q-O-P* genes with a deletion in *cbiK*. All strains for siroheme/cysteine complementation are derived from strain ER171: *E. coli cysG* deleted strain (302Δa) containing pACYC184/*lacI*$^q$. Growth on minimal media occurs in all constructs when supplemented with cysteine.

rium between a trimeric and monomeric state within the conditions of the crystallographic experiments.

A search for structurally related proteins with the DALI server (*26*) pairs CbiK with the *B. subtilis* PPIX ferrochelatase (PDB code: 1ak1) with a *Z*-score of 18.6 and an rms deviation of 3.1 Å over 233 residues (Figure 2B,C). Several additional type-I periplasmic-type binding proteins show a reduced but significant similarity. Previously, the periplasmic-type binding proteins were classified as type-I containing six parallel strands (213456) and type-II containing a mixed sheet of five strands with strand 5 antiparallel (21354). Together, CbiK and PPIX ferrochelatase create a new subclass, type-III, containing four parallel strands (2134).

There are a number of significant structural differences between CbiK and PPIX ferrochelatase. In CbiK, helix α1′ is inserted between β2-α2, while in ferrochelatase, two helices, α1′ and α1″, are inserted between a different pair of secondary structural units, β1-α1. The extra helices occupy approximately the same 3-dimensional space in both structures (Figure 2A,B,C), and are situated near a cleft in the center of the molecule identified as the active site (see below). The superposition between CbiK and ferrochelatase suggests a potential intramolecular domain motion, as superposition of each domain independently results in a better fit than for that of the whole molecule. Aligning the C-terminal domains (upper) of both enzymes places the N-terminal domain of ferrochelatase 17° in front and to the left of the N-terminal domain of CbiK (Figure 2D). The combination of the extra helical subdomain and the tighter packing of the N- and C-terminal domains in PPIX ferrochelatase results in a relatively inaccessible active site, whereas the active site cleft in CbiK is open.

Although a direct sequence comparison of CbiK and PPIX ferrochelatase suggests there is no similarity, a structure-based sequence alignment indicates a modest 11% identity (Figure 3A,B). Of the 32 residues that are identical between the 2 structures (black), only 14 are completely conserved within the extended CbiK family (blue). Of these, nine are involved in maintaining the overall structure: Lys2, Leu6, Gly11, Ile57, Ala65, Pro178, Pro199, Ala204, and Gly288. The remaining conserved residues surround the active site and/or participate in the catalytic mechanism; Thr12, Tyr14, His145, Glu175, Asp211.

*Active Site.* The active site of CbiK lies in a deep rectangular cleft (~8 × 8 × 20 Å) between the two domains (Figures 2−4). The identification of this area as the active site is supported by analogy to PPIX ferrochelatase (*15*) as well as by the presence of a number of conserved residues. Each protein requires a unique set of residues around the active site to bind selectively its tetrapyrrole-derived substrate. In CbiK, the active site contains a mixture of charged and neutral residues to interact with the precorrin-2 substrate, which contains eight carboxylic acid moieties attached to the tetrapyrrole framework through nonpolar side chains. PPIX ferrochelatase has to bind a substrate that contains a more delocalized macrocyclic ring system and a much more hydrophobic arrangement of peripheral side chains.

In CbiK, His145 lies at the end of β5 and is predicted to function as the principal ligand for the cobalt ion substrate, homologous to His183 in PPIX ferrochelatase (Figure 3A,B). In PPIX ferrochelatase, the neighboring residue on α7, Glu264, is thought to act as the general base in the abstraction of protons from the pyrrole nitrogens. In contrast, the equivalent position on α7 of CbiK is occupied by a second histidine, His207. These two imidazole groups of CbiK are positioned such that a poorly ordered water molecule sits between them on the surface of the protein (Figure 4A). It is likely that the second histidine contributes to the cobalt binding site and effects the ion selectivity of the enzyme although it could also act as a general base. The proposed substrate binding site for PPIX ferrochelatase places the porphyrin ring between the general base and the metal ion ligand within the C-terminal domain (*15*). In contrast, the accessibility, shape, and orientation of the CbiK active site cleft suggest that the substrate binds between the two domains (Figure 4B). This revised orientation places the substrate such that both proton abstraction and metal ion

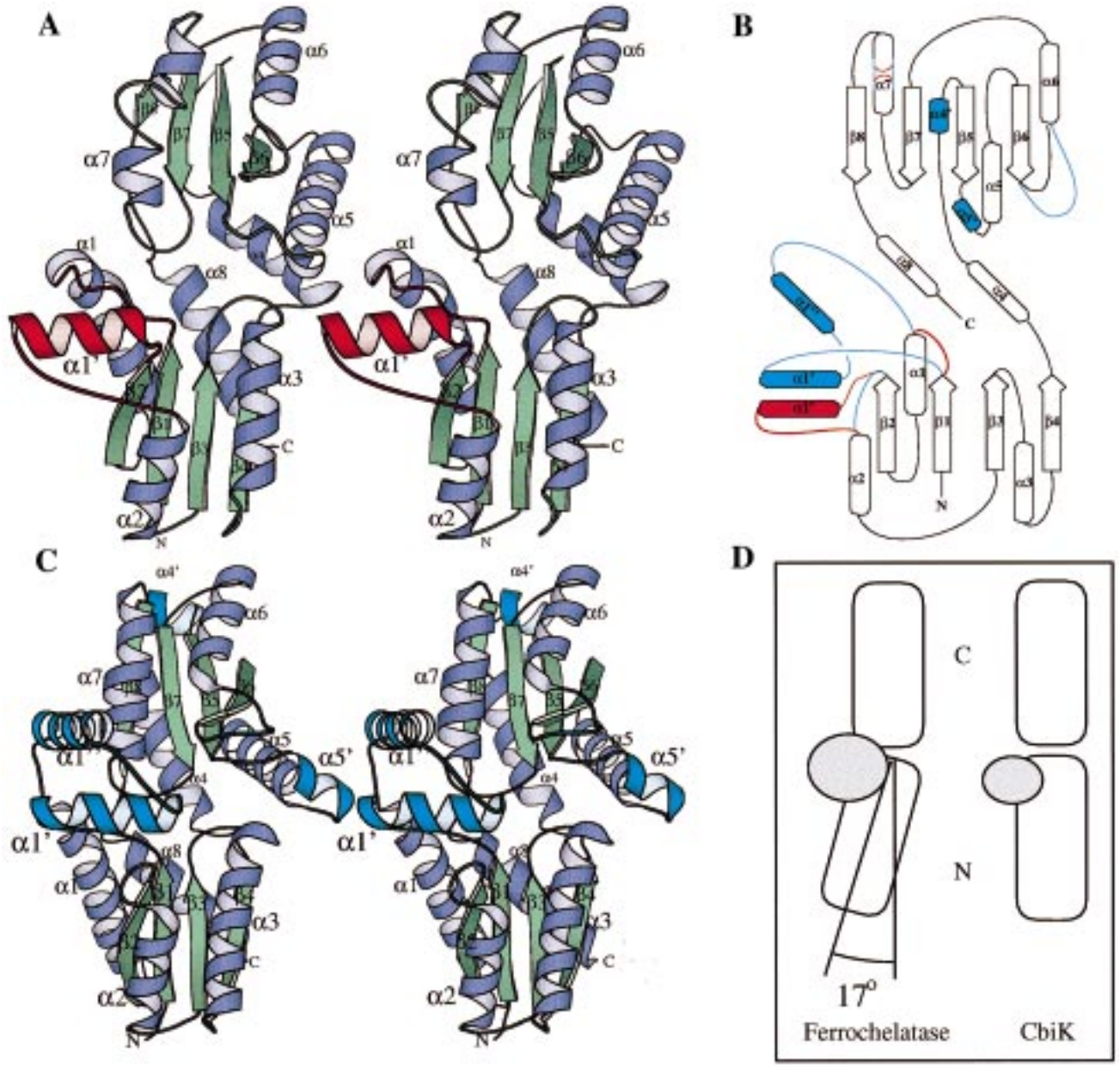


FIGURE 2: Structure of CbiK. (A) Stereo diagram of a secondary structure schematic of the anaerobic cobalt chelatase, CbiK (*35*, *36*). (B) topology diagram reflecting not only the fold of CbiK but also regions of divergence with the PPIX ferrochelatase structure (CbiK shown in red). (C) Stereo diagram of the structure of PPIX ferrochelatase, in which structurally distinct regions are colored cyan. (D) Alignment of the C-terminal (upper) domains of CbiK and ferrochelatase results in a 17° shift of the N-terminal domain of ferrochelatase upward and to the left. This results in a relatively inaccessible active site in PPIX ferrochelatase, whereas the active site in CbiK consists of an open and distinct cleft between the two domains.

insertion will take place on the same side of the tetrapyrrole ring, in accordance with earlier predictions (*27*).

However, removal of protons from precorrin-2 could also be achieved by CbiK binding the substrate in a puckered conformation such that protons are removed from two opposing pyrrole nitrogens, one from either face of the molecule. This would require a general base positioned on the opposite side of the cleft. In CbiK, conserved Glu89 lies on the opposite side of the active site, a position related to His207 by the pseudo-2-fold internal symmetry (gene-duplication event) of the molecule (Figure 4A), and could potentially function as a general base. The essential nature of Glu89 was tested through site-directed mutagenesis (see below).

The presence of conserved acidic residues in the C-terminal domain and several nonconserved positively charged residues in the N-terminal domain effect a remarkable charge polarity on the electrostatic surface of CbiK (Figure 5A). Charge−charge interactions may participate in the attraction and binding of both the negatively charged precorrin-2 substrate containing eight carboxylic acid side chains and the positively charged cobalt(II) ion. In the C-terminal domain, several negatively charged residues are located just above the active site; Glu175 and Asp211 lie at the end of $\beta$5 and within $\alpha$7, respectively. The position of these residues is such that they could participate in metal ion ligation in conjunction with His145. However, since a hydrogen bond links the side chains of Asp211 and His207, the role of

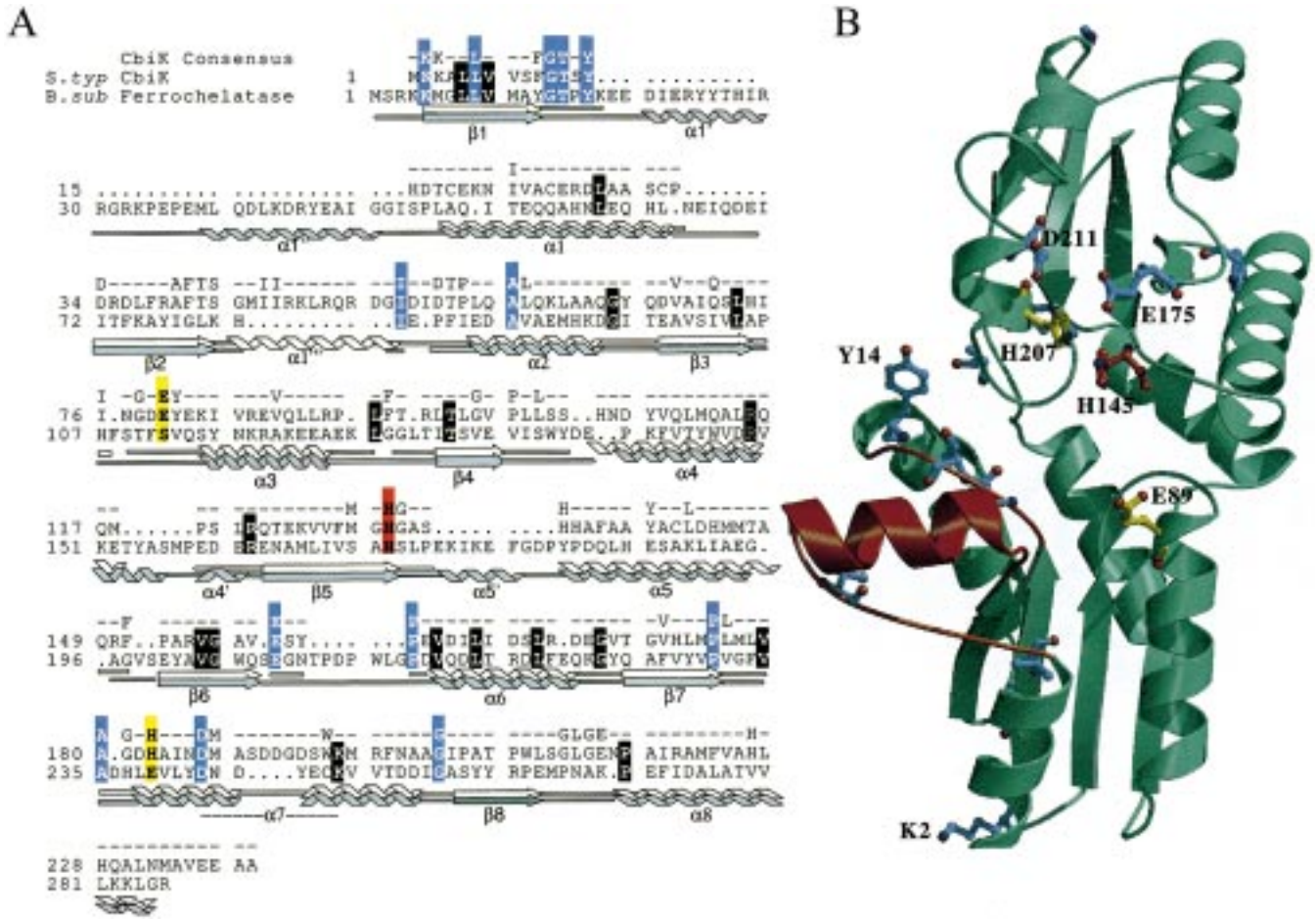


FIGURE 3: Conservation of amino acid sequence and 3D structure. (A) Amino acid sequence alignment between *S. typhimurium* CbiK and *B. subtilis* PPIX ferrochelatase. The consensus sequence for known CbiKs is provided on the top line, and the structure of both enzymes is annotated underneath the relevant sequence: CbiK, white; ferrochelatase, gray. Residues conserved in all three sequences are shown in purple while residues conserved only within the two structures are shown in black. The principal cobalt protein ligand is His145 shown in red, but contributions toward substrate binding are also thought to come from His207 and Glu89, highlighted in yellow. (B) The structure of CbiK featuring conserved residues annotated according to the colors described in (A).

Asp211 may be to position the histidine rather than directly ligate the cobalt ion.

The functional significance of the surface charge disparity is supported by the comparison with PPIX ferrochelatase where the C-terminal domain is negatively charged and the N-terminal domain is relatively nonpolar (Figure 5B). These domains would interact with the positive ferrous(II) ion and a relatively hydrophobic protoporphyrin. Since neither cobalt nor iron is expected to be unligated inside a cell, the charge polarity might also help recruit cellular proteins involved in storage and transfer of metal ions.

The $\beta$1-$\alpha$1 loop contains conserved residues Thr12 and Tyr14. Thr12 points into the active site and is in a position to form both hydrogen bonds and hydrophobic interactions to the acetic and propionic acid side chains (Figure 4A). In CbiK, Tyr14 is positioned on $\alpha$1 and forms hydrogen bonds across the domain interface, suggesting a role in, or modulation of, domain movements in response to substrate binding. The corresponding residue in ferrochelatase, Tyr17, plays a similar role albeit in the opposite direction, since it lies in the loop leading to the helical subdomain and anchors it back to the N-terminal domain though a hydrogen bond.

Opposite His145 is the aromatic ring of Phe10 (Tyr8 in PPIX ferrochelatase) which is positioned to provide a hydrophobic interaction with the pyrrole rings of the substrate (Figure 4A,B,C). The ring may provide a backstop to the incoming metal ion. A similar configuration is present in methionine synthase (*28*) (PDB code: 1BMT), where a histidine (His759) residue ligates the cobalt ion on one side of cobalamin, and positions a phenylalanine (Phe708) on the opposite side of the ring (Figure 4C). In methionine synthase, Phe708 is further removed from the cobalamin to avoid clashes with the active methyl group ligated to the cobalt. Alignments with methylmalonyl-CoA mutase (*29*) (PDB code: 4REQ) do not reveal a similar configuration since the substrate binding site occupies the space opposite the histidine ligand. Two tyrosine residues do flank the cobalamin ring opposite the histidine, but they are offset from the cobalt ion, toward the edge of the ring structure (Figure 4C).

*Site-Directed Mutagenesis.* Site-directed mutagenesis was employed to probe the role played by key amino acids identified from the structure. The residues selected were His145 and His207, which appear to form a cobalt binding site; Glu89, a candidate for the general base from the opposite side of the substrate cleft; and Glu175 and Asp211, potential participants in a secondary cobalt binding site or in direct conjunction with the histidine residues. All residues were modified to alanine as described under Materials and Methods, and the double mutant H145A/H207A was also

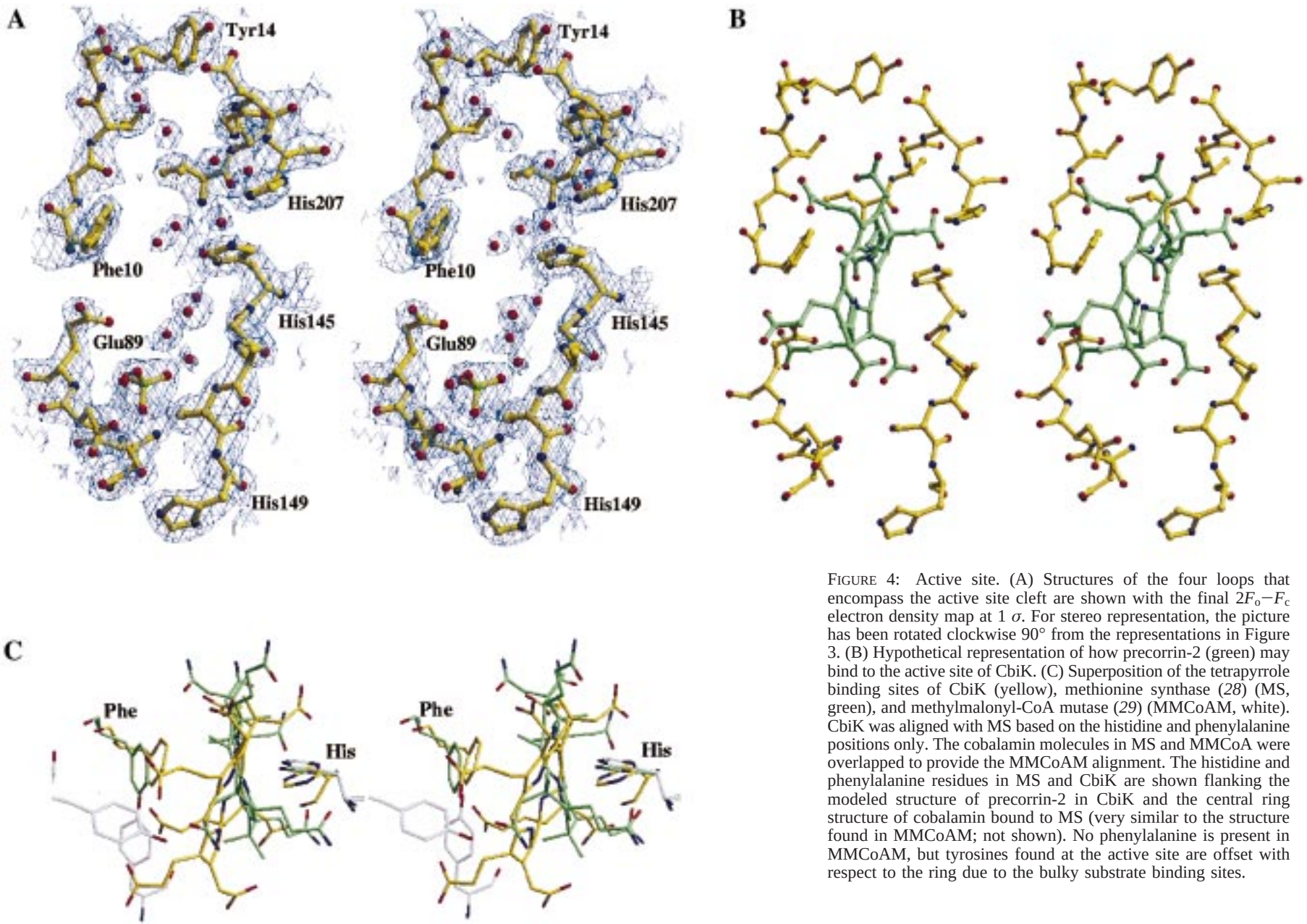


FIGURE 4: Active site. (A) Structures of the four loops that encompass the active site cleft are shown with the final $2F_o-F_c$ electron density map at 1 $\sigma$. For stereo representation, the picture has been rotated clockwise 90° from the representations in Figure 3. (B) Hypothetical representation of how precorrin-2 (green) may bind to the active site of CbiK. (C) Superposition of the tetrapyrrole binding sites of CbiK (yellow), methionine synthase (*28*) (MS, green), and methylmalonyl-CoA mutase (*29*) (MMCoAM, white). CbiK was aligned with MS based on the histidine and phenylalanine positions only. The cobalamin molecules in MS and MMCoA were overlapped to provide the MMCoAM alignment. The histidine and phenylalanine residues in MS and CbiK are shown flanking the modeled structure of precorrin-2 in CbiK and the central ring structure of cobalamin bound to MS (very similar to the structure found in MMCoAM; not shown). No phenylalanine is present in MMCoAM, but tyrosines found at the active site are offset with respect to the ring due to the bulky substrate binding sites.

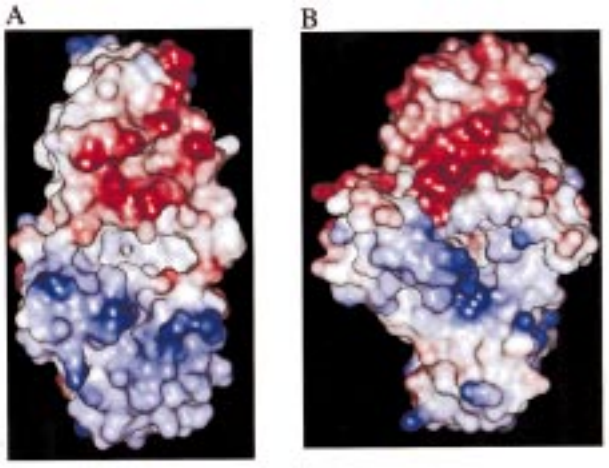


FIGURE 5: Molecular surface colored by electrostatic potential. The molecular surface of CbiK (A) and PPIX ferrochelatase (B) color-coded according to the electrostatic potential of the surface residues: blue, positive; red, negative (MOLVIEWER-OGL). The strong negative potential of the upper domain is involved in cobalt-(II) and ferrous binding, and the more positively charged lower domain may interact with the negatively charged precorrin-2 and protoporphyrin IX substrates.

constructed. The effects of these alterations on the activity of CbiK were analyzed in three ways.

First, CbiK can act as a precorrin-2 ferrochelatase since, in conjunction with a uroporphyrinogen III methyltransferase, such as CobA, it can complement an *E. coli cysG* deletion strain, 302Δa (*7*). However, this complementation is easily inhibited in the presence of low concentrations of exogenous cobalt, a result that can be explained on the basis of a higher specificity for cobalt over ferrous ions (Table 4). The mutants harboring $CbiK^{E89A}$, $CbiK^{D211A}$, and $CbiK^{E175A}$ were found to have the same phenotype as the wild-type enzyme in that they complemented *E. coli* 302Δa only in the absence of exogenous cobalt (Table 4). However, the histidine mutants displayed an altered phenotype. In the absence of exogenous cobalt, $CbiK^{H207A}$ efficiently complements the *cysG* deletion strain. In contrast, $CbiK^{H145A}$ was barely able to overcome the cysteine auxotrophy, indicating this mutant has greatly reduced ferrochelatase activity. Surprisingly, both histidine mutants were also able to complement the *cysG* deletion strain in the presence of exogenous cobalt, suggesting that they now have an altered, lower specificity for cobalt. Finally, the double histidine mutant ($CbiK^{H145A/H207A}$) was unable to restore growth to the *cysG* deletion strain, and is therefore inactive as a precorrin-2 ferrochelatase in siroheme synthesis.

Second, the activity of the CbiK variants was also examined through the spectrophotometric analysis of accumulated intermediate(s) in strains overproducing the mutant chelatase and a uroporphyrinogen methylase (CobA). The visible spectrum of the tetrapyrrole-derived compounds from the strain overproducing CobA alone is largely consistent with the presence of trimethylpyrrocorphin (maxima absorption ~350 nm), although an absorption maximum at ~380 nm suggests that there is some sirohydrochlorin within the sample (Figure 6a) (*30, 31*). The trimethylpyrrocorphin appears as a result of an extra, nonphysiological, methylation of precorrin-2 at position 12 by the presence of an excess of CobA. Addition of exogenous cobalt to the growth medium does not modify the CobA spectrum. When *cbiK* is coexpressed with *cobA*, the spectrum, in the absence of cobalt, indicates that sirohydrochlorin accumulates (Figure 6b), though an additional peak at 406 nm may be due to the formation of sirohydrochlorin monolactone. In the presence of cobalt, two new maxima at 414 and 595 nm appear, and the spectrum is very similar to that previously reported for cobalt sirohydrochlorin (*30*).

Spectroscopic analysis of the strains that contain CobA and the various mutant CbiK constructs suggests that the CbiK variants can be separated into three categories (Figure 6c−h). The first includes $CbiK^{E89A}$ and $CbiK^{D211A}$, which produce similar spectra to wild-type CbiK. The second category contains $CbiK^{E175A}$, which synthesizes mainly trimethylpyrrocorphin but, in the presence of cobalt, can also produce cobalt sirohydrochlorin at a reduced level. The third includes the various histidine alterations, $CbiK^{H145A}$, $CbiK^{H207A}$, and $CbiK^{H145A/H207A}$, which give spectra similar to the strain containing only CobA, producing mainly trimethylpyrrocorphin. In summary, the three histidine variants are unable to catalyze the efficient chelation of cobalt to generate cobalt sirohydrochlorin.

Finally, the *in vivo* activity of the CbiK variants was tested by determining their ability to restore cobyric acid synthesis to a recombinant *E. coli, cbiK* deficient strain (ER185). The production of cobyric acid by several of the mutant CbiKs, including $cbiK^{E89A}$, $cbiK^{D211A}$ and $cbiK^{E175A}$, was similar to that of the wild-type CbiK both in the presence and in the absence of exogenous cobalt (Table 4). The mutant $cbiK^{H207A}$ also produces wild-type levels of cobyric acid in the presence of cobalt, but in the absence of cobalt the amount of measurable corrin product was significantly reduced. The most debilitating single mutation with respect to corrin biosynthesis is the His145Ala variant. This enzyme restores very little cobyric acid synthesis in the presence of cobalt and does not produce any cobyric acid in the absence of exogenous cobalt. Finally, the $cbiK^{H145A/H207A}$ double mutant does not promote any cobyric acid production under any of the conditions, confirming that the encoded protein is completely inactive.

*Mechanistic Implications.* Site-directed mutagenesis confirms that the principal protein ligand for cobalt is His145, and that His207 and Glu175 may also participate. The two histidine residues also appear to affect the selectivity for cobalt over iron. Asp211 and Glu89 appear to have no direct effect on the overall catalytic efficiency, though the experiments are not sensitive enough to detect altered tetrapyrrole affinities and moderate reductions in catalytic activity.

Glu89 can be ruled out as a general base in the catalytic mechanism. Proton abstraction from the tetrapyrrole could alternatively take place through either His145 or His207. While distortion of the porphyrin ring upon metal binding has been observed in ferrochelatases (*27, 32*), there is no physical evidence to date for such a mechanism during the precorrin-2 chelation of cobalt. However, an induced conformational change of precorrin-2 upon binding to CbiK, favoring proton abstraction and metal insertion, is the most likely mechanism. The synthesis of the metallo−precorrin-2 complex would then presumably trigger product release.

*Structural Evolution.* Proteins evolve due to both functional and substrate-specific pressures. Recent structure determinations have provided examples of both structurally related proteins which are divergent in function and function-

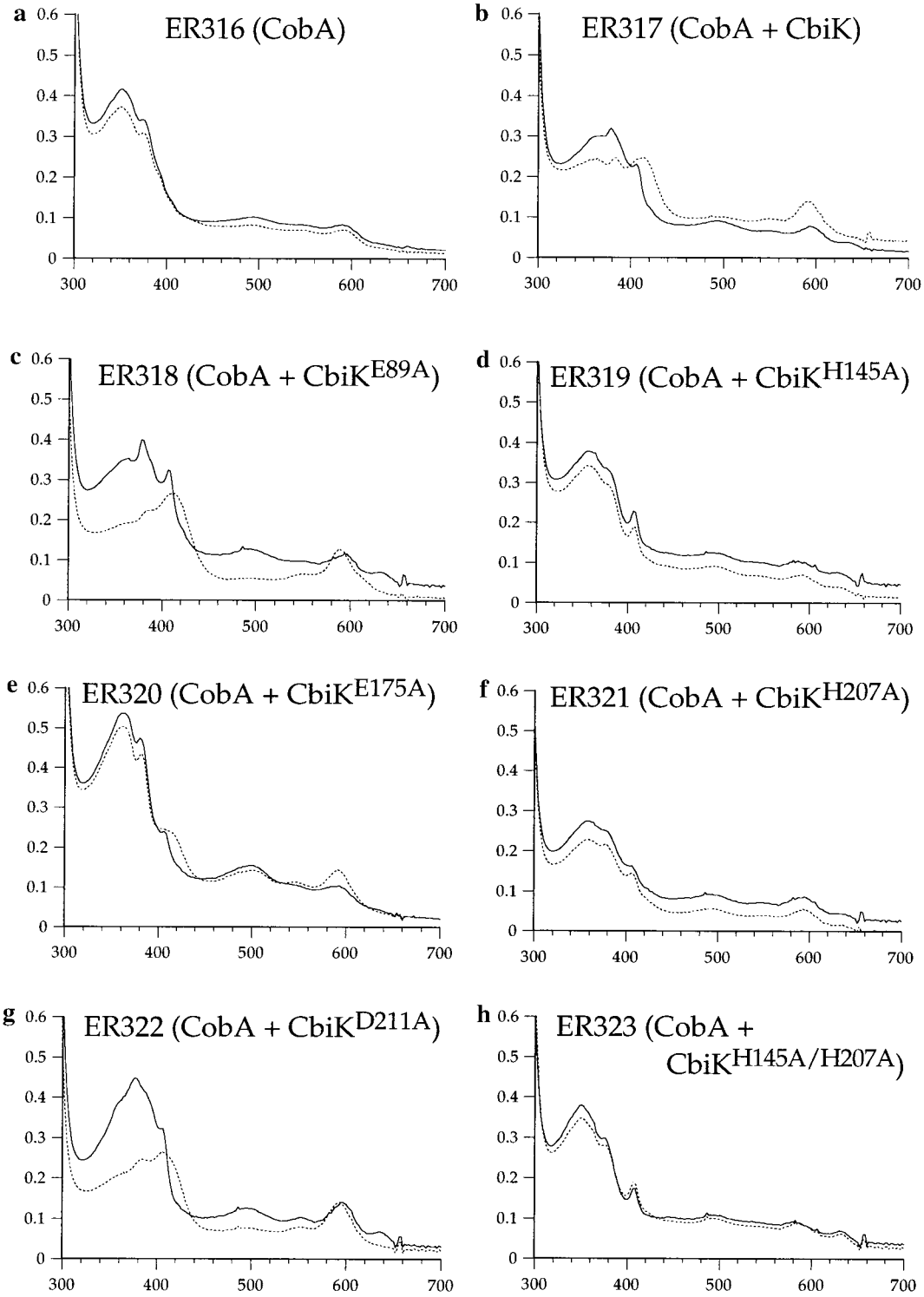


FIGURE 6: Spectra of the tetrapyrrole-derrived compounds in wild-type and mutant strains. Visible spectra between 300 and 700 nm of purified tetrapyrrolic extracts from bacterial strains expressing wild-type or mutant constructs of CbiK along with CobA. (a) CobA control, (b) wild-type CbiK, (c) CbiK$^{H145A}$, (d) CbiK$^{E89A}$, (e) CbiK$^{A175A}$, (f) CbiK$^{H207A}$, (g) CbiK$^{D211A}$, (h) CbiK$^{H145A/H207A}$. Peaks at 350, 380, and 414/590 nm are consistent with the accumulation of trimethylpyrrocorphin, sirohydrochlorin, and cobalt sirohydrochlorin, respectively. Plain lines indicate spectra from extracts of cells grown in the absence of cobalt while dashed lines represent spectra of cell extracts grown in the presence of cobalt.

ally related proteins which are divergent in structure (*33*). The comparison of CbiK and ferrochelatase provides an example of structural homology with significant sequence diversity linking two related biosynthetic pathways through a common functionality. CbiK represents the more ancient structure since cobalamin is regarded as the matriarch of this prosthetic group family. Moreover, the CbiK structure is simpler and does not contain as many inserts as the PPIX ferrochelatase.

The cobalt chelatase in *B. megaterium*, CbiX, has no significant sequence similarity to the known protoporphyrin or precorrin-2 chelatases. By analogy to the similarity between the PPIX ferrochelatase and CbiK structures, we might still predict that CbiX will have a similar tertiary

structure. A more extensive alignment search between CbiK and CbiX reveals weak identity between the C-terminal domain of CbiK and the N-terminus of CbiX (*9*). Many of the relevant histidine and acidic residues align with few sequence gaps. This suggests that the fold of CbiX could approximate that of the C-terminal domain of CbiK and will either leave the other side of the active site uncovered or, more likely, be enclosed by the remaining C-terminal fragment. Moreover, CbiK also displays some similarity with *Yersinia enterocolitica* HemR (30% identity in 103 amino acid overlap), an iron-regulated outer membrane hemin receptor protein (*34*), suggesting that this particular fold may have been adopted for purposes other than chelation. The slight similarity may reflect either a tetrapyrrole binding site or a ferrous binding epitope.

Two further independent sequences for monomeric chelatases are known: the ferrochelation domain of CysG and a yeast enzyme, Met8p[6]. Neither enzyme shows any homology to CbiK or PPIX ferrochelatase, and they may represent novel, evolutionary divergent, chelatase structures. The aerobic cobalt chelatase and magnesium chelatase appear to be a completely distinct family (*2*), reflecting perhaps a later evolutionary appearance associated with the advent of phototrophic life.

## REFERENCES


1. Warren, M. J., and Scott, A. I. (1990) *Trends Biochem. Sci. 12*, 486−491.
2. Jensen, P. E., Gibson, L. C. D., Henningsen, K. W., and Hunter, C. N. (1996) *J. Biol. Chem. 271*, 1662−1667.
3. Debussche, L., Couder, M., Thibaut, D., Cameron, B., Crouzet, J., and Blanche, F. (1992) *J. Bacteriol. 174*, 7445−7451.
4. Dailey, H. A. (1997) *J. Biol. Inorg. Chem. 2*, 411−417.
5. Warren, M. J., Bolt, E. L., Roessner, C. A., Scott, A. I., Spencer, J. B., and Woodcock, S. C. (1994) *Biochem. J. 382*, 837−844.
6. Raux, E., McVeight, T., Peters, S. E., Leustek, T., and Warren, M. J. (1999) *Biochem. J. 338*, 701−708.
7. Raux, E., Thermes, C., Heathcote, P., Rambach, A., and Warren, M. J. (1997) *J. Bacteriol. 330*, 121−129.
8. Beck, R., Raux, E., Thermes, C., Rambach, A., and Warren, M. J. (1997) *Biochem. Soc. Trans. 25*, 77S.
9. Raux, E. (1999) University of London.
10. Lavallee, D. K. (1988) in *Mechanistic principles of enzymatic activity* (Liebmann, J. F., and Greenberg, A., Eds. pp 279−314, VCH, New York.
11. Dailey, H. A., Jones, C. S., and Karr, S. W. (1989) *Biochim. Biophys. Acta 999*, 7−11.
12. Dailey, H. A. (1996) in *Mechanisms of metallocenter assembly* (Hausinger, R. P., Eichhorn, G. L., and Marzelli, L. G., Eds. pp 77−98, VCH, New York.
13. Gora, M., Grzybowska, E., Rytka, J., and Labbe-Bois, R. (1996) *J. Biol. Chem. 271*, 11810−11816.
14. Kohno, H., Okuda, M., Furukawa, T., Tokunaga, R., and Taketani, S. (1994) *Biochim. Biophys. Acta 1209*, 95−100.
15. Al-Karadaghi, S., Hansson, M., Nikonov, S., Jonsson, B., and Hederstedt, L. (1997) *Structure 5*, 1501−1510.
16. Raux, E., Schubert, H. L., Woodcock, S. C., Wilson, K. S., and Warren, M. J. (1998) *Eur. J. Biochem. 254*, 341−346.
17. Wilson, A. J. C. (1942) *Nature 150*, 151−152.
18. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
19. Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
20. The CCP4 Suite: Programs for Protein Crystallography (1994) *Acta Crystallogr. D50*, 760−763.
21. Cowtan, K. (1994) *Jt. CCP4 ESF-EACBM Newslett. Protein Crystallogr. 31*, 34−48.
22. Oldfield, T. J. (1996) in *Macromolecular Refinement* (Dodson, E. J., Moore, M. H., Ralph, A., and Bailey, S., Eds.) pp 67−74, VHC, New York.
23. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D53*, 240−255.
24. Raux, E., Lanois, A., Levillayer, F., Warren, M. J., Brody, E., Rambach, A., and Thermes, C. (1996) *J. Bacteriol. 178*, 753−767.
25. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) *J. Mol. Biol. 247*, 536−540.
26. Holm, L., and Sander, C. (1995) *Trends Biochem Sci. 20*, 478−480.
27. Blackwood, M. E., Jr., Rush, T. S., III, Medlock, A., Dailey, H. A., and Spiro, T. G. (1997) *J. Am. Chem. Soc. 119*, 12170−12174.
28. Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G., and Ludwig, M. L. (1994) How a protein binds $B_{12}$: A 3.0 Å X-ray structure of $B_{12}$-binding domains of methionine synthase. *Science 266*, 1669−1674.
29. Mancia, F., Keep, N. H., Nakagawa, A., Leadlay, P. F., McSweeney, S., Rasmussen, B., Bosecke, P., Diat, O., and Evans, P. R. (1996) How coenzyme $B_{12}$ radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 Å resolution. *Structure 4*, 339−350.
30. Battersby, A. R., and Sheng, Z. C. (1982) *J. Chem. Soc., Chem Commun. D*, 1393−1394.
31. Bykhovskii, V. Y., Santander, P. J., Stupperikh, E., Zaitseva, N. I., Pusheva, M. A., Detkova, E. N., Valyushok, D. S., and Scott, A. I. (1996) *Appl. Biochem. Microbiol. 32*, 170−178.
32. Blackwood, M. E., Jr., Rush, T. S., III, Romesberg, F., Schultz, P. G., and Spiro, T. G. (1998) *Biochemistry 37*, 779−782.
33. Martin, A. C. R., Orengo, C. A., Hutchinson, E. G., Jones, S., Karmirantzou, M., Laskowski, R. A., Mitchell, J. B. O., Taroni, C., and Thornton, J. M. (1998) *Structure 6*, 875−884.
34. Stojiljkovic, I., and Hantke, K. (1992) *EMBO J. 11*, 4359−4367.
35. Merrit, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869−873.
36. Esnouf, R. M. (1997) *J. Mol. Graphics 15*, 133−138.


BI9906773