Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Structural Biology 153 (2006) 307–311
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

Crystallization note

The crystal structure of putative precorrin isomerase CbiC in cobalamin biosynthesis

Yanyan Xue [a,b], Zhiyi Wei [a,b], Xu Li [b], Weimin Gong [a,b,*]

[a] National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, PR China
[b] School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230026, PR China

Received 28 July 2005; received in revised form 24 November 2005; accepted 29 November 2005
Available online 27 December 2005

Abstract

The leptospira *cbiC* encodes the enzyme catalyzing the methyl rearrangement reaction of the cobalamin biosynthesis pathway. The protein has been cloned and overexpressed as a His-tagged recombinant protein in *Escherichia coli*. The crystal structures have been solved in two crystal forms ($P4_22_12$ and $P3_121$) diffracting to 3.0 and 2.3 Å resolution, respectively. The structures are similar to the precorrin-8x methyl mutase (CobH), an enzyme of the aerobic pathway to vitamin B12.
© 2005 Elsevier Inc. All rights reserved.

Keywords: *Leptospira interrognans*; Leptospirosis; Vitamin B12; Precorrin isomerase; Crystal structure

1. Introduction

Cobalamin (vitamin B12) is one of the most structurally complex protein cofactors of the nature. Most prokaryotic organisms as well as animals (including humans) and protists have enzymes that require cobalamin as cofactor, whereas plants and fungi are thought not to use it. Among the cobalamin-utilizing organisms, only some bacterial and archaeal species are able to synthesize cobalamin de novo (Rodionov et al., 2003). Studies over the past decade have demonstrated that there are at least two quite distinct routes for cobalamin synthesis, representing oxygen-dependent (aerobic) and oxygen-independent (anaerobic) pathways (Scott and Roessner, 2002). The two pathways differ in their timing of cobalt insertion and the requirement for molecular oxygen (Raux et al., 1999; Schubert et al., 1999).

All the intermediates between uro'gen III and cobinamide are known for the aerobic pathway, in contrast, few intermediates on the anaerobic pathway are known. The genes encoding the aerobic pathway enzymes are prefixed Cob, whereas the genes encoding the anaerobic pathway are prefixed Cbi (Roper et al., 2000). Many of the aerobic, Cob, enzymes share a high degree of similarity with the anaerobic, Cbi, suggesting that although independent, the two pathways are broadly similar (Roth et al., 1993).

Leptospira interrognans is most common in tropical areas and infects animals. Infection can cause mild symptoms or more serious disease, known as leptospirosis. Here, we report the crystal structure of the *L. interrognans* putative CbiC (CbiC_LEPIN) with two different crystal forms, $P4_22_12$ and $P3_121$. CbiC, a cobalt-precorrin 8 isomerase, catalyzes cobalt-precorrin 8 to cobyrinic acid by methyl rearrangement in anaerobic pathway (Roessner et al., 1992; Roth et al., 1993). The specific homology of CbiC in aerobic pathway is CobH, which shows high similarity to CbiC despite the divergence between the two pathways. Structure analysis suggests that CbiC_LEPIN are highly similar with CbiC/CobH from other sources in fold and has small differences in surrounding of cobalt-precorrin 8 binding. The CbiC_LEPIN structure will provide clues for understanding the function of CbiC/CobH family in the biosynthesis pathway.

* Corresponding author.
E-mail address: wgong@sun5.ibp.ac.cn (W. Gong).

1047-8477/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2005.11.011

2. Materials and methods

2.1. Protein expression and purification

The gene *cbiC* from *L. interrogans* was cloned into a pET-22b (Novagen) expression plasmid to yield pET-His-cbiC. The plasmid was transformed into *Escherichia coli* strain BL21 (DE3) for protein expression. Five milliliter aliquots of an overnight culture were subcultured into 400 ml fresh Luria–Bertani medium (10 g Bacto tryptone, 5 g yeast extract, and 10 g NaCl per liter of solution) containing ampicillin (50 μg/ml) and allowed to grow at 310 K until $OD_{600} = 0.6$. At this stage isopropyl-β-D-thiogalactoside was added to a final concentration of 0.2 mM and the cells were grown for a further 20 h at 289 K. Then cells were harvested by centrifugation at 8000 rpm for 10 min. The cell pellet was resuspended in binding buffer (20 mM Tris/HCl, pH 7.5, 0.2 M NaCl) and sonicated using a JY92-II sonic dismembrator. The solution was centrifuged again (16000 rpm, 4.5 min) and the supernatant was applied in a Ni-affinity column equilibrated with binding buffer. The column was then washed with 10 column volumes of binding buffer, 6 column volumes of wash buffer (50–100 mM imidazole, 0.2 mM NaCl, and 20 mM Tris/HCl, pH 7.5), and finally the protein was eluted in 6 column volumes of elution buffer (200 mM imidazole, 0.2 mM NaCl, and 20 mM Tris/HCl, pH 7.5). The fractions were checked by SDS–PAGE and those containing CbiC_LEPIN protein were concentrated by ultrafiltration equipment. The final protein concentration was measured with a Bio-Rad Protein Assay kit (Bio-Rad Pacific, USA).

2.2. Crystallization and preliminary data collection

Preliminary crystallization conditions were screened using Crystal Screen I and II (Hampton Research, CA, USA) with the handing-drop vapour-diffusion technique at 277 K. Drops containing equal volumes (1 μl) of protein (15 mg/ml) and reservoir solution were equilibrated against 400 μl reservoir solution. Two conditions ([10%(v/v) PEG6000, 2.0 M sodium chloride] and [20%(v/v) ethanol, 0.1 M Tris, pH 8.5]) produced CbiC_LEPIN protein crystals with $P4_22_12$ and $P3_121$ crystal forms, respectively. More suitable crystals for diffracting experiment were obtained in [15%(v/v) ethanol and 0.1 M Tris, pH 8.0] after optimizing the second condition. Preliminary diffraction data were collected on a MAR Research image-plate system with a local X-ray source at 100 K with 1.5418 Å wavelength. All data were processed and scaled with the DENZO and SCALEPACK (Otwinowski and Minor, 1997). The processing statistics of the two different crystal forms are summarized in Table 1.

2.3. Structure determination

The crystal structure of CbiC_LEPIN with the $P4_22_12$ crystal form was solved using Molrep (Vagin and Teplyakov, 1997). The monomer structure of Precorrin-8x Methyl Mutase from *Thermus thermophilus* (PDB code: 1V9C) was used as a search model for the molecular replacement. The model consisting of a dimer was refined in 20–3.0 Å resolution range by using CNS (Brunger et al., 1998) with maximum-likelihood amplitude targets and manually adjusted and rebuilt of the model using the program O (Jones et al., 1991) with $2F_o - F_c$ and $F_o - F_c$ electron-density maps as references. Water molecules were added to the model and individual atomic B factor was refined at latter stage of refinement. NCS restraints were applied through all stages of refinement. The final R_{cryst} and R_{free} factors are 20.4 and 25.1%, respectively. A monomer structure of this refined model was used as a search model to find solutions for the $P3_121$ crystal form with one molecule per asymmetric unit. The refinement strategy

Table 1
Data collection and structure refinement statistics of two forms crystal of CbiC_LEPIN

Data collection		
Space group	$P4_22_12$	$P3_121$
Unit cell parameters	$a = b = 113.08$, $c = 114.50$	$a = b = 55.89$, $c = 142.55$
Molecules per asymmetric unit	2	1
Resolution range (Å)	20–3.0 (3.07–3.00)	50–2.3 (2.35–2.30)
No. of total reflections	195344	132312
No. of unique reflections	15336	10964
Completeness (%)	99.6 (95.6)	90.8 (99.9)
I/σ	15.73 (4.57)	23.29 (7.38)
R_{merge} (%)[a]	0.112 (0.392)	0.061 (0.239)
Structure refinement		
Resolution (Å)	20–3.0 (3.19–3.0)	50–2.3 (2.44–2.3)
R_{cryst}/R_{free} (%)[b]	20.4 (30.2)/25.1 (38.1)	22.4 (20.5)/28.4 (26.8)
No. of reflections		
Working set	13439	9632
Test set	1520	1100
Rmsd from ideal values		
Bond length (Å)	0.006	0.005
Bond angels (°)	1.2	1.1
Average *B*-factor (Å^2)		
Main chain	37.8/32.7[c]	33.4
Side chain	38.3/34.0[c]	35.8
No. of atoms		
Protein	3284	1654
Solvent	71 (including 8 chloride ions)	75
Ramachandran plot		
Most favored regions (%)	85.1	88.1
Additionally allowed (%)	14.4	11.9
Generously allowed (%)	0.5	

Numbers in parentheses are for the highest resolution shell.

[a] $R_{merge} = \sum|I_i - I_m|/\sum I_i$, where I_i is the intensity of the measured reflection and I_m is the mean intensity of all symmetry-related reflections.

[b] $R_{cryst} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure factors. $R_{free} = \sum_T||F_{obs}| - |F_{calc}||/\sum_T|F_{obs}|$, where T is a test data set of about 10% of the total reflections randomly chosen and set aside prior to refinement.

[c] Values for the two different monomers (A and B) respectively in an asymmetric unit.

of the $P3_121$ structure is similar to the $P4_22_12$ structure with the final R_{cryst} and R_{free} factors 22.4 and 28.4%, respectively, in 50–2.3 Å resolution range. The stereochemical qualities of the final models of two crystal forms were checked by PROCHECK (Laskowski et al., 1993), and the final refinement statistics and geometry are listed in Table 1. The relatively high R_{free} factor value of the $P3_121$ structure is mainly dues to existing obvious ice rings during data collection.

3. Results and discussion

3.1. Protein preparation and crystallization

The recombinant protein of CbiC_LEPIN was overexpressed at 310 K at first, but the protein was insoluble. Various expression conditions were tried to increase solubility of the protein. Finally, the recombinant protein was expressed solubly in higher yield under the following conditions: 289 K, 0.2 mM IPTG, 20 h induction in strain BL21 (DE3). The protein was purified using the nickel-affinity column and the sample was monitored by SDS–PAGE after purification and concentration. The purified His-tagged putative CbiC protein was used for crystallization. After an initial screen, two crystal forms of the protein were observed. The polyhedral crystal form was only obtained a 7–8 Å resolution data at first. After optimization and using a cryoprotectant solution (10% glycerol) before flash freezing the crystal to liquid nitrogen, the resolution was improved greatly. The maximum resolution was 2.3 Å and the space group was hexagonal $P3_121$, with unit-cell parameters $a = b = 55.89$ Å, $c = 142.55$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$. The rod-shaped crystal form was obtained a 3.0 Å resolution data set. The space group was tetragonal $P4_22_12$, with unit-cell parameters $a = b = 113.08$ Å, $c = 114.50$ Å, $\alpha = \beta = \gamma = 90°$.

3.2. Structure overview

The final models contain 215 residues in one molecule of $P3_121$ crystal form and 434 residues (217 residues each molecule) in two molecules (forming a dimer) of $P4_22_12$. N-terminal residues, Met1–Gln5 (in $P3_121$) and Met1–Asp3 (in $P4_22_12$), the C-terminal residues, Glu222–Arg223 (in $P3_121$) and Gly221–Arg223 (in $P4_22_12$), as well as the His_tag were not visible in electron-density maps. The three monomer structures (one in $P3_121$ and two in $P4_22_12$) are almost identical with the average root-mean-square deviation (rmsd) about 0.46 Å (for 215 Cα between $P3_121$ and any one of $P4_22_12$ monomer structures) or about 0.26 Å (for 217 Cα between two $P4_22_12$ monomer structures). The CbiC_LEPIN structure contains 11 α-helices and a six-stranded mixed β-sheet in which the five strands (β2–β6) are parallel and the edge strand β1 run anti-parallel to β6. The β-sheet is surrounded by four α-helices in one side and seven in another (Fig. 1A).

CbiC_LEPIN forms a dimer in both of the two crystal forms with a crystallographic (in $P3_121$ crystal) or a non-crystallographic (in $P4_22_12$ crystal) 2-fold symmetry between the two monomers (Fig. 1B). The dimer is formed by both hydrophobic and hydrophilic interactions (including some salt bridges.) The dimerization form of CbiC_LEPIN is in agreement with other CbiC/CobH structures. It suggests that dimerization is necessary for having biological function in CbiC/CobH family. Two identical cavities formed by two CbiC/CobH molecules in a dimer are also essential for their substrates binding (Shipman et al., 2001).

3.3. Comparison with other CbiC/CobH

Sequence and structural alignment reveal that CbiC_LEPIN is fairly similar to other CbiC/CobH (Figs. 2A and B). Farther structural analyses show that CbiC/CobH

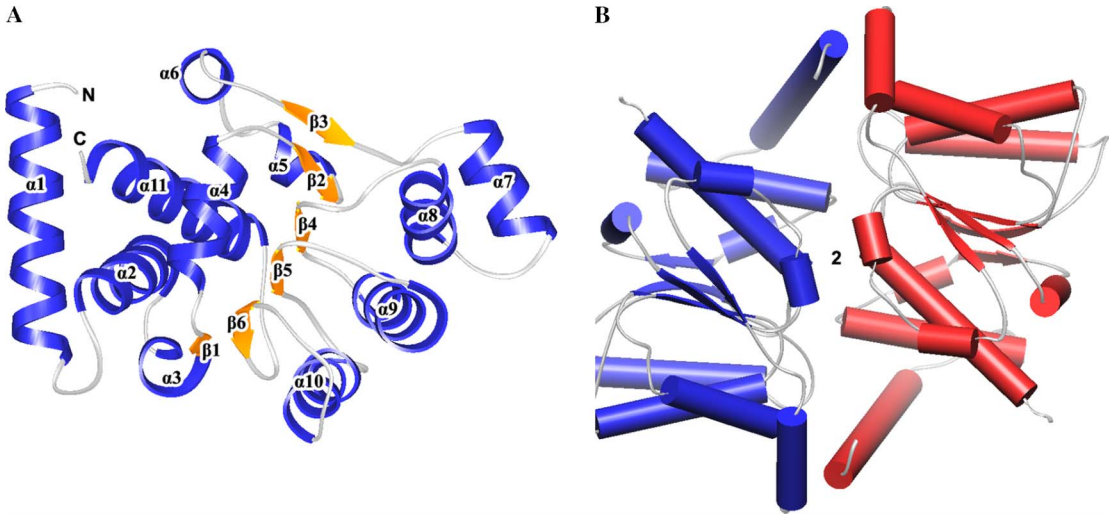

Fig. 1. The ribbons diagram of the monomer and the dimer structure of CbiC_LEPIN. (A) The protein is color-coded by secondary structure types. (B) The two monomers are related by a crystallographic (in $P3_121$ structure) or a non-crystallographic (in $P4_22_12$ structure) 2-fold screw axis where is showed as "2." This figure and Fig. 2B were prepared with RIBBONS (Carson, 1997).

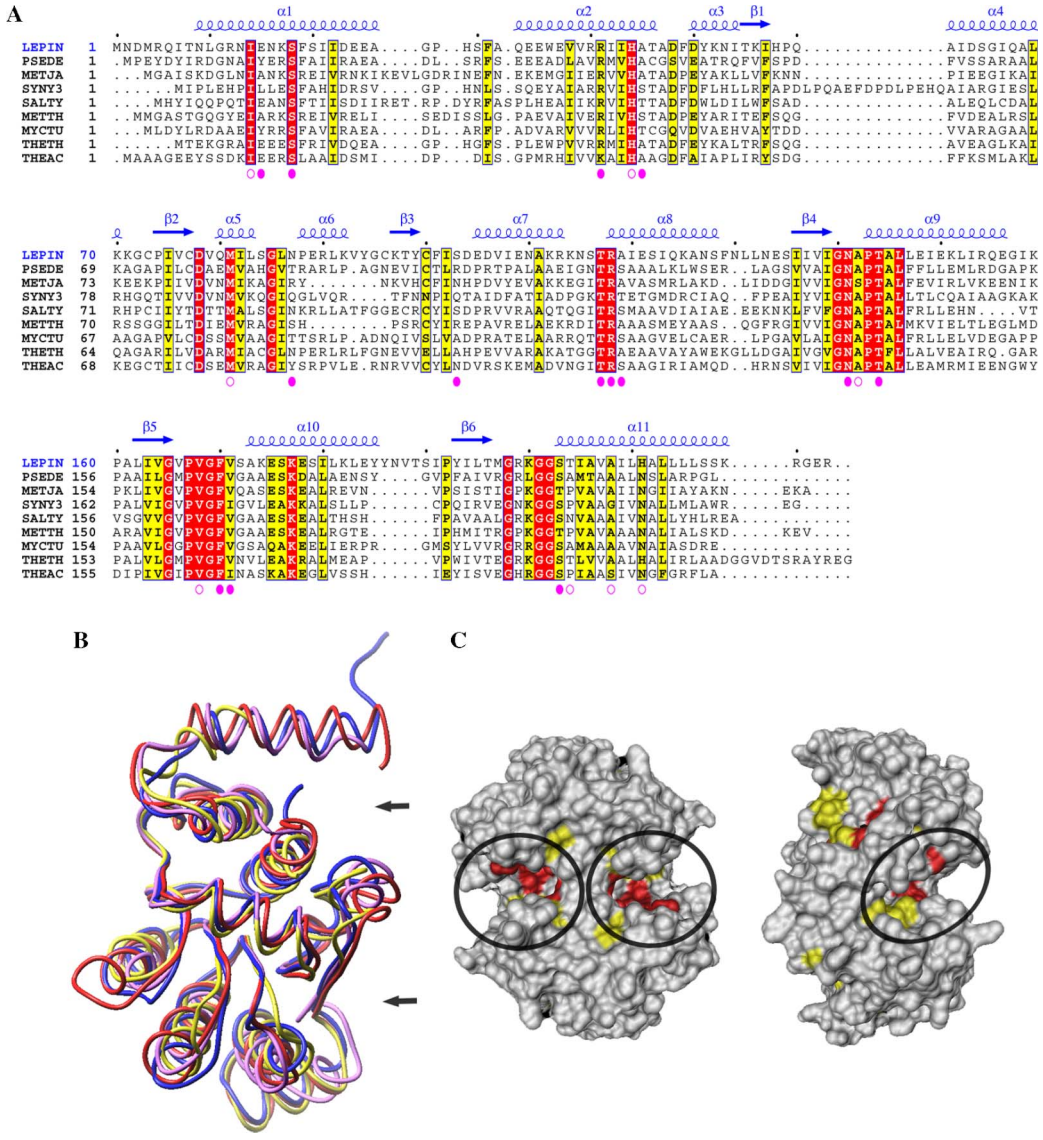

Fig. 2. (A) Sequence alignment of CbiC_LEPIN and other CbiC/CobH protein. The sequences of CbiC/CobH from *Leptospira interrognans* (LEPIN), *Pseudomonas denitrificans* (PSEDE), *Methanococcus jannaschii* (METJA), *Synechocystis* sp. strain PCC 6803 (SYNY3), *Salmonella typhimurium* (SALTY), *Methanobacterium thermoautotrophicum* (METTH), *Mycobacterium tuberculosis* (MYCTU), *Thermus thermophilus* (THETH), and *Thermoplasma acidophilum* (THEAC) were aligned using ClustalW (Notredame et al., 2000). The last two protein sequences come from PDB, no related information of these two proteins can be obtain from other protein sequence databases; thus, the two protein are putative CbiC/CobH protein. The alignment was edited with referenced to structural alignment results to avoid gaps inside the conserved secondary structural elements. The secondary structure of CbiC_LEPIN, which is defined by the analysis of the structure using DSSP program (Kabsch and Sander, 1983), is indicated above the alignment. Residues in the alignment that are identical are shown in red boxes; those that are similar are shown in yellow boxes. The residues that are involved in hydrophobic interactions or hydrophilic interactions with HBA in structure of CobH_PSEDE (PDB id: 1I1H) are marked by solid circles or hollow circles, respectively. This figure was prepared with ESPript (Gouet et al., 1999). (B) The superposition of the structure of CbiC_LEPIN (red), CobH_PSEDE (blue), and other CbiC/CobH from *Thermus thermophilus* (pink, PDB id: 1V9C) and *Thermoplasma acidophilum* (yellow, PDB id: 1OU0). The structural alignment was performed using O (Jones et al., 1991). The two black arrows denote that the locations of halves of the two identical cavities, respectively, in monomer structures. (C) Surface distributions of conserved residues in CbiC_LEPIN. The two identical cavities are marked by circles on the left picture. The right picture is a 90° rotated view around a vertical axis in which only one cavity can be seen.

family members also have high similarity in shape of the two identical cavities (Fig. 2B) and centralized distributions of conserved residues on and around the surface of the cavities (Figs. 2A and C). Recently, the complex structure of *Pseudomonas denitrificans* CobH (CobH_PSEDE) binding with hydrogenobyrinic acid (HBA), which is the product of the reaction catalyzed by CobH, has been solved (Shipman et al., 2001). This structure shows that the cavities supply the binding surface for the substrates. And comparing with the complex structure, conformation of residues on the cavities of the CobH_PSEDE native structure does not show obvious changes. Thus, although without substrates binding, the similarity of the cavities in CbiC/CobH family suggests that cobalt-precorrin 8,

which is the substrate of CbiC, is highly similar to the substrate of CobH, precorrin 8.

Although the residues that composed the cavities are mostly conserved, there are still some non-conserved residues distribute on the surface of the cavities. These non-conserved residues might respond for different substrates (cobalt-precorrin 8 or precorrin 8) binding. In addition, N-terminal helices functioning to cap the active site of cavity in CbiC/CobH structures are probably disordered in a precorrin-8x methylmutase related protein from *Thermoplasma acidophilum* (Fig. 2B), which is a putative CbiC/CobH family member.

There is little biochemical characterization of CbiC_LEPIN and consequently the information that can be derived from the structure reported here is limited. Whether binding mode of cobalt-precorrin 8 is highly similar to that of precorrin 8 and what the cobalt-precorrin 8 is, need to be further investigated.

Acknowledgments

This work was supported by National Foundation of Talent Youth (Grant No. 30225015), The 973 programs (Grant Nos. 2004CB720000 and 2004CB520801), the Key Important Project and other projects from the National Natural Science Foundation of China (Grant Nos. 10490913, 30121001, and 30130080) and Chinese Academy of Sciences (KSCX2-SW-224). We thank Prof. Peng Liu and Yuhui Dong in Institute of High Energy Physics, and Yi Han in Institute of Biophysics for diffraction data collection. The atomic coordinates and structure factors of the CbiC structure of the two crystal forms ($P3_121$ and $P4_22_12$) have been deposited in the Protein Data Bank as 2AFR and 2AFV, respectively.

References

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., Warren, G.L., 1998. Crystallography & NMR system (CNS): a new software suite for macromolecular structure determination. Acta. Crystallogr. D 54, 905–921.

Carson, M., 1997. Ribbons. Methods Enzymol. 277, 493–505.

Gouet, P., Courcelle, E., Stuart, D.I., Metoz, F., 1999. ESPript: multiple sequence alignments in PostScript. Bioinformatics 15, 305–308.

Jones, T.A., Zou, J.Y., Cowan, S.W., Kjeldgaard, M., 1991. Improved methods for building protein models in electron density maps and the location of error in these models. Acta Crystallogr. A 47, 110–119.

Kabsch, W., Sander, C., 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22, 2577–2637.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26, 283–291.

Notredame, C., Higgins, D.G., Heringa, J., 2000. A novel algorithm for multiple sequence alignment. J. Mol. Biol. 302, 205–217.

Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. In: Carter, C.W., Sweet, R.M. (Eds.), Methods in Enzymology 276, Macromolecular Crystallography, part A, Academic Press, New York, pp. 307–326.

Raux, E., Schubert, H.L., Roper, J.M., Wilson, K.S., Warren, M.J., 1999. Vitamin B12: insights into biosynthesis's mount improbable. Bioorg. Chem. 27, 100–118.

Rodionov, D.A., Vitreschak, A.G., Mironov, A.A., Gelfand, M.S., 2003. Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. J. Biol. Chem. 278, 41148–41159.

Roessner, C.A., Warren, M.J., Santander, P.J., Atshaves, B.P., Ozaki, S., Stolowich, N.J., Iida, K., Scott, A.I., 1992. Expression of 9 *Salmonella typhimurium* enzymes for cobinamide synthesis. Identification of the 11-methyl and 20-methyl transferases of corrin biosynthesis. FEBS Lett. 301, 73–78.

Roper, J.M., Raux, E., Brindley, A.A., Schubert, H.L., Gharbia, S.E., Shah, H.N., Warren, M.J., 2000. The enigma of cobalamin (vitamin B12) biosynthesis in *Porphyromonas gingivalis*. J. Biol. Chem. 275, 40316–40323.

Roth, J.R., Lawrence, J.G., Rubenfield, M., Kieffer-Higgins, S., Church, G.M., 1993. Characterization of the cobalamin (vitamin B12) biosynthetic genes of *Salmonella typhimurium*. J. Bacteriol. 175, 3303–3316.

Schubert, H.L., Raux, E., Wilson, K.S., Warren, M.J., 1999. Common chelatase design in the branched tetrapyrrole pathways of heme and anaerobic cobalamin synthesis. Biochemistry 38, 10660–10669.

Scott, A.I., Roessner, C.A., 2002. Biosynthesis of cobalamin (vitamin B(12)). Biochem. Soc. Trans. 30, 613–620.

Shipman, L.W., Li, D., Roessner, C.A., Scott, A.I., Sacchettini, J.C., 2001. Crystal structure of precorrin-8x methyl mutase. Structure 9, 587–596.

Vagin, A., Teplyakov, A., 1997. MOLREP: an automated program for molecular replacement. J. Appl. Cryst. 30, 1022–1025.