Cobalamin (vitamin B₁₂) biosynthesis: functional characterization of the *Bacillus megaterium* *cbi* genes required to convert uroporphyrinogen III into cobyrinic acid *a,c*-diamide

Evelyn RAUX^{*1}, Anne LANOIS[†], Alain RAMBACH[‡], Martin J. WARREN^{*} and Claude THERMES[†]

^{*}Department of Molecular Genetics, Institute of Ophthalmology, UCL, 11–43 Bath Street, London EC1V 9EL, U.K., [†]Centre de Génétique Moléculaire, Laboratoire associé à l'Université Pierre et Marie Curie, Centre National de la Recherche Scientifique, 91198 Gif sur Yvette, France, and [‡]CHROMagar, 267 rue Lecourbe, 75015 Paris, France

The function of individual genes of the *Bacillus megaterium* *cobI* operon genes in cobalamin (vitamin B₁₂) biosynthesis was investigated by their ability to complement defined *Salmonella typhimurium* *cob* mutants. This strategy confirmed the role of *cbiA*, *-D*, *-F*, *-J*, *-L* and *cysG*⁴. Furthermore the operon as a whole was used to restore corrin biosynthesis in *Escherichia coli*, which, although closely related to *S. typhimurium*, does not possess the CobI pathway. When the *B. megaterium* *cob* operon was cloned into a plasmid and transformed into an *E. coli* strain containing the *S. typhimurium* *cbiP*, it conferred upon the host

strain the ability to make the cobyrinic acid *de novo*. However, cobyrinic acid synthesis was observed only when the strain was grown anaerobically. Derivatives of the corrin-producing *E. coli* strain were constructed in which genes of the *B. megaterium* *cob* operon had been inactivated. These strains were used to demonstrate that, whereas *B. megaterium* *cbiD*, *-G* and *-X* are essential for cobyrinic acid synthesis, the *cbiW* and *-Y* genes could be deleted without detriment to cobyrinic acid production in *E. coli*.

INTRODUCTION

About 25 enzymes are required to transform uroporphyrinogen III into cobalamin [1,2]. In the obligate aerobic *Pseudomonas denitrificans*, the process by which uroporphyrinogen III is transformed into adenosylcobyrinic acid has been completely elucidated and has been shown to require 16 enzymes (CobA, *-I*, *-G*, *-J*, *-M*, *-F*, *-L*, *-K*, *-H*, *-B*, *-N*, *-S*, *-T*, reductase, *-O* and *-Q*) (outlined in Scheme 1) [2,3]. The first of these enzymes, uroporphyrinogen III transmethylease, converts uroporphyrinogen III into precorrin-2 by two *S*-adenosyl-L-methionine-dependent transmethylation at positions 2 and 7. A further methylation at C-20 catalysed by the precorrin-2 transmethylease from *P. denitrificans* produces precorrin-3a, which is the substrate for the molecular oxygen-requiring CobG. This enzyme enhances the lactonization of precorrin-3a, yielding precorrin-3b. Ring contraction of precorrin-3b is promoted during the subsequent transmethylation of position C-17, a reaction catalysed by CobJ that yields precorrin-4. A further four *S*-adenosyl-L-methionine-dependent transmethylation (CobM, *-F* and *-L*), deacylation (CobF), macrocyclic ring reduction (CobK), decarboxylation (CobL), methyl rearrangement (CobH) and amidations (CobB) produces hydrogenobyrinic acid *a,c*-diamide. Cobalt chelation by an enzyme complex formed from CobN, *-S* and *-T*, cobalt reduction, amidations of the remaining peripheral side chains by CobQ and adenosylation, catalysed by CobO, produces adenosylcobyrinic acid (reviewed in [2–6]).

Recently it has become apparent that there are major differences between the aerobic cobalamin pathway of *P. denitrificans* and the anaerobic corrin biosynthesis of *Salmonella typhimurium*. These differences relate mostly to the timing of

cobalt insertion and to the ring-contraction mechanism (Scheme 1) [4,7,8]. The three *P. denitrificans* genes required for cobalt insertion, *cobS*, *-N* and *-T*, are not found in *S. typhimurium* [1]; conversely *S. typhimurium* *cbiK*, which seems to be required for cobaltochelation of precorrin-2 [9], is not found in *P. denitrificans* [2]. An *Escherichia coli* strain harbouring all the *S. typhimurium* *cbi* genes on a plasmid produced only the early cobalamin precursors precorrin-2 and -3 when the growth medium was not supplemented with cobalt [8]. In the presence of cobalt, full corrin production was restored with no trace of precorrin-2 or -3. Moreover, the two genes involved in ring contraction in *P. denitrificans*, *cobF* and *cobG*, are not found in *S. typhimurium*, whereas the *S. typhimurium* *cobI* operon contains *cbiD* and *cbiG*, which have no homologue in *P. denitrificans* [8]. The function of these latter two gene products has not yet been determined but they are known to be essential for cobalamin biosynthesis [8]. In effect, only one intermediate between cobalt-precorrin-3 and cobyrinic acid has been identified, cobalt-precorrin-4 [10]. A model for two different cobalamin pathways has been proposed, involving the *S. typhimurium* or *Propionibacterium shermanii* anaerobic pathway, and the *P. denitrificans* aerobic pathway [4,7,8].

A *Bacillus megaterium* *cob* operon has recently been identified and fully sequenced (the preceding paper [11]). It was found to contain 14 genes, 11 of which shared some degree of identity with cobalamin biosynthetic genes found in *S. typhimurium* and *P. denitrificans*. To investigate the necessity of these genes for cobalamin biosynthesis, complementation studies were undertaken with defined *S. typhimurium* *cob* mutants. Although *S. typhimurium* and *E. coli* are closely related enteric species, the latter is unable to synthesize cobalamins *de novo* as it does not

Abbreviation used: IPTG, isopropyl β -D-thiogalactoside.

¹ To whom correspondence should be addressed (e-mail email e.raux@ucl.ac.uk).

Scheme 1 Outline biosynthesis of adenosylcobyrinic acid from uroporphyrinogen III

The diagram highlights the differences between the anaerobic and aerobic pathways. In the aerobic pathway molecular oxygen is required as a substrate and cobalt is added comparatively late. In the anaerobic pathway cobalt is added at the level of precorrin-2. Abbreviation: uro'gen III, uroporphyrinogen III.

possess any of the genes required to convert precorrin-2 into cobinamide [12] (i.e. it is lacking the CbiI pathway). The capacity to biosynthesize corrins was restored by cloning the whole

B. megaterium cob operon into *E. coli*, which consequently allowed a more detailed functional characterization of some of the *B. megaterium cob* operon genes.

MATERIALS AND METHODS

Chemicals and reagents

Cyanocobalamin, cobinamide and most chemicals were purchased from Sigma. Cobyric acid was synthesized from cyanocobalamin as described previously [13]. Restriction enzymes and isopropyl β -D-thiogalactoside (IPTG) were purchased from Appligene; calf intestinal alkaline phosphatase and T4 DNA

ligase were from Promega; pKK223.3 and Klenow fragment of DNA polymerase I were purchased from Pharmacia.

Bacterial strains, media and growth conditions

All the *E. coli* strains containing plasmids carrying fragments of the *B. megaterium cob* operon were derived from AR3730 (LE392 rendered met⁺ [8]). The genotypes of all the strains are listed in

Table 1 List of strains and plasmids used in this study

<i>E. coli</i> strain	Genotype and/or phenotype	Description	Reference or source
AR3730	LE392 met ⁺		[8]
AR8925	LE392 met ⁺ (pAR8849)(pAR8911)	Contains <i>S.t cbiP</i> and all <i>B.m cbi</i> genes, a mutation in <i>cbiX</i>	This work
AR8930	LE392 met ⁺ (pAR8849)(pAR8916)	Contains <i>S.t cbiP</i> and all <i>B.m cbi</i> genes except the 2nd part of <i>cbiH₆₀</i>	This work
AR8937	LE392 met ⁺ (pAR8849)(pAR8877)	Contains <i>S.t cbiP</i> and all <i>B.m cbi</i> genes	This work
AR8967	LE392 met ⁺ (pAR8962)(pAR8911)	Contains <i>S.t cbiP</i> and <i>B.m cbiX</i> and all <i>B.m cbi</i> genes, a mutation in <i>cbiX</i>	This work
ER144	LE392 met ⁺ (pAR8849)(pAR8846)	Contains <i>S.t cbiP</i> and all <i>B.m cbi</i> genes except <i>cbiW</i>	This work
ER164	LE392 met ⁺ (pAR8849)(pAR8909)	Contains <i>S.t cbiP</i> and all <i>B.m cbi</i> genes except <i>cbiD</i>	This work
ER165	LE392 met ⁺ (pAR8849)(pAR8921)	Contains <i>S.t cbiP</i> and all <i>B.m cbi</i> genes, a mutation in <i>cbiG</i>	This work
ER167	LE392 met ⁺ (pER156)(pAR8909)	Contains <i>S.t cbiP</i> and <i>B.m cbiD</i> and all <i>B.m cbi</i> genes, a mutation in <i>cbiD</i>	This work
ER169	LE392 met ⁺ (pER160)(pAR8921)	Contains <i>S.t cbiP</i> and <i>B.m cbiG</i> and all <i>B.m cbi</i> genes, a mutation in <i>cbiG</i>	This work
ER176	LE392 met ⁺ (pER146)(pAR8877)	Contains <i>S.t cbiP</i> and <i>P.d cobA</i> and all <i>B.m cbi</i> genes	This work
ER177	LE392 met ⁺ (pER146)(pAR8980)	Contains <i>S.t cbiP</i> and <i>P.d cobA</i> and all <i>B.m cbi</i> genes except <i>cysG^A</i> , <i>cbiY</i> and <i>btuR</i>	This work
ER241	LE392 met ⁺ (pAR8849)(pAR8766)	Contains <i>S.t cbiP</i> & all <i>B.m cbi</i> genes (no <i>tac</i> promoter)	This work
LE392	F ⁻ e14 ⁻ (mcrA ⁻) <i>hsdR514</i> (<i>r_K</i> ⁻ <i>m_K</i> ⁻) <i>supE44 supF58 lacY1</i> or Δ (<i>lacZY</i>)6 galK2 galT22 metB1 trpR55		[14]
Plasmid	Inserted genes	Description	Reference or source
pAR8086	<i>E. coli lacI^q</i>		[8]
pAR8531	<i>S.t cbiP</i>		[8]
pAR8754	<i>S.t cbiA</i> - <i>B.m cbiX</i> - <i>J-C-D-ET</i>		[11]
pAR8766	<i>B.m cbiW</i> - <i>H₆₀</i> - <i>X-J-C-D-ET-L-F-G-A-cysG^A</i> - <i>cbiY</i> - <i>btuR</i> -><-ORF1-ORF2		[11]
pAR8799	<i>B.m cysG^A</i>	<i>NaeI</i> 12313- <i>SspI</i> 13162 fragment of pAR8766 cloned into the <i>EcoRI</i> site of pKK223.3	This work
pAR8811	<i>B.m cbiH₆₀</i> - <i>X-J</i>	1914 to 5834 <i>MluI</i> fragment cloned into the <i>SmaI</i> site of pKK223.3	This work
pAR8846	<i>B.m cbiH₆₀</i> - <i>X-J-C-D-ET-L-F-G-A-cysG^A</i> - <i>cbiY</i> - <i>btuR</i> -><-ORF1-ORF2	<i>NcoI</i> 2971- <i>SalI</i> 16432 fragment of pAR8766 cloned into <i>NcoI</i> / <i>SalI</i> of pAR8811; this plasmid does not contain <i>cbiW</i>	This work
pAR8849	<i>S.t cbiP</i>	<i>SphI</i> - <i>Scal</i> fragment of pAR8531 (<i>Ptac</i> - <i>S.t cbiP</i>) cloned into <i>SphI</i> / <i>NruI</i> sites of pAR8086	This work
pAR8877	<i>B.m cbiW</i> - <i>H₆₀</i> - <i>X-J-C-D-ET-L-F-G-A-cysG^A</i> - <i>cbiY</i> - <i>btuR</i> -><-ORF1-ORF2	<i>BglI</i> 1218- <i>SalI</i> 16432 fragment of pAR8766 cloned into <i>SmaI</i> / <i>SalI</i> sites of pKK223.3	This work
pAR8882	<i>B.m cbiX</i>	<i>SspI</i> 3642- <i>SnaBI</i> 5041 fragment of pAR8766 cloned into the <i>SmaI</i> site of pKK223.3	This work
pAR8909	<i>B.m cbiW</i> - <i>H₆₀</i> - <i>X-J-C-D^M</i> - <i>ET-L-F-G-A-cysG^A</i> - <i>cbiY</i> - <i>btuR</i> -><-ORF1-ORF2	Mutation at the <i>AccI</i> 6412 site into the <i>cbiD</i> gene of pAR8877 (TAG at position 6444)	This work
pAR8911	<i>B.m cbiW</i> - <i>H₆₀</i> - <i>X^M</i> - <i>J-C-D-ET-L-F-G-A-cysG^A</i> - <i>cbiY</i> - <i>btuR</i> -><-ORF1-ORF2	Mutation at the <i>AccI</i> 4088 site into the <i>cbiX</i> gene of pAR8877 (TAA at position 4135)	This work
pAR8916	<i>B.m cbiW</i> - <i>H₆₀</i> - <i>X^M</i> - <i>J-C-D-ET-L-F-G-A-cysG^A</i> - <i>cbiY</i> - <i>btuR</i> -><-ORF1-ORF2	Mutation at the <i>AccI</i> 3241 site into the 2nd part of <i>cbiH₆₀</i> gene of pAR8877 (TGA at position 3295)	This work
pAR8921	<i>B.m cbiW</i> - <i>H₆₀</i> - <i>X-J-C-D-ET-L-F-G^M</i> - <i>A-cysG^A</i> - <i>cbiY</i> - <i>btuR</i> -><-ORF1-ORF2	Mutation at the <i>Clal</i> 10217 site into the <i>cbiG</i> gene of pAR8877 (TAA at position 10411)	This work
pAR8962	<i>S.t cbiP</i> - <i>B.m cbiX</i>	<i>EcoRI</i> - <i>Scal</i> digest of pAR8882 cloned into <i>SalI</i> site of pAR8849	This work
pAR8980	<i>B.m cbiW</i> - <i>H₆₀</i> - <i>X-J-C-D-ET-L-F-G-A</i>	<i>EcoRI</i> fragment of pAR8877 (from 1218 to 12734) cloned into <i>EcoRI</i> site of pKK223.3	This work
pCR395	<i>P.d cobA</i>		[16]
pER110	<i>B.m cbiL</i>	<i>EcoRV</i> 7968- <i>BglII</i> 9235 fragment of pAR8766 cloned into <i>SmaI</i> - <i>BamHI</i> sites of pKK223.3	This work
pER146	<i>S.t cbiP</i> - <i>P.d cobA</i>	<i>EcoRI</i> - <i>SalI</i> fragment of pCR395 cloned into <i>SalI</i> site of pAR8849	This work
pER156	<i>S.t cbiP</i> - <i>B.m cbiD</i>	5999 to 7755 <i>PvuII</i> fragment of pAR8766 cloned into <i>SalI</i> site of pAR8849	This work
pER160	<i>S.t cbiP</i> - <i>B.m cbiG</i>	<i>PstI</i> 9471- <i>KpnI</i> 11061 fragment of pAR8766 cloned into <i>SalI</i> site of pAR8849	This work
pER180	<i>B.m cbiD</i>	5999 to 7755 <i>PvuII</i> fragment of pAR8766 cloned into <i>SmaI</i> site of pKK223.3	This work
pER187	<i>B.m cbiA</i>	<i>NdeI</i> 11007- <i>MluI</i> 12661 of pAR8766 cloned into <i>SmaI</i> site of pKK223.3	This work
pKK223.3		Overexpression vector derived from pBR322 with <i>tac</i> promoter; ampicillin resistant	Pharmacia

Table 1. All *S. typhimurium* strains used in this study have been discussed previously [8] except AR8647, which is described in Table 1. Other strains were grown routinely in Luria–Bertani broth or Luria–Bertani agar at 37 °C, with 100 mg/l ampicillin and 30 mg/l chloramphenicol as required. To assess cobalamin production the strains were grown in minimal medium supplemented with 0.1 g/l yeast extract. Complementation of *S. typhimurium cob* mutants was determined as described previously [8].

Subcloning the *cob* operon under the control of *tac* promoter

Cloning procedures were performed as described in [14]. The *B. megaterium cobI* operon was isolated originally on pAR8766, a promoterless pKK233-3 derivative. The operon was placed under the regulation of a *tac* promoter by cloning the *Bgl*I–*Sal*I fragment of pAR8766 into the *Sma*I and *Sal*I sites of pKK223.3, giving rise to the plasmid pAR8877. A second plasmid was also constructed that did not contain the first gene of this operon (*cbiW*). This was achieved by restricting pAR8811 (pKK223.3 carrying the *Mlu*I fragment from pAR8766; Table 1) with *Nco*I and *Sal*I and ligating the *Nco*I and *Sal*I restriction fragments of pAR8766 into it. The resulting plasmid, pAR8846, contained all the *B. megaterium cobI* operon genes except the first open reading frame (*cbiW*).

Mutations or deletion introduced into *cbiH*₆₀, *cbiD*, *cbiG* and *cbiX*

A mutation in *cbiH*₆₀ was generated as follows: pAR8877 was partly restricted with *Acc*I, and blunt ends were generated by the action of the Klenow fragment. The linearized plasmid was isolated and purified by agarose gel electrophoresis and was subsequently religated by the action of DNA ligase. This procedure generated a 2 bp insertion into *cbiH*₆₀ (codons 2161–3781) resulting in a frameshift and a new stop codon at position 3295 and producing plasmid pAR8916. A similar procedure was used to introduce mutations into the *Acc*I sites found in *cbiX* and *cbiD*, and with the *Cla*I site of *cbiG*. The 2 bp introduced into *cbiX* (codons 3787–4705) led to a stop codon at position 4135 which was harboured in plasmid pAR8911. The 2 bp introduced in *cbiD* (6129–7230) led to a stop codon at position 6444 and the mutated plasmid was named pAR8909. The 2 bp insertion into *cbiG* (9892–11014) led to a stop codon at position 10411 that was harboured in plasmid pAR8921 (Table 1).

Deletion of the *cysG*^A-*cbiY*-*btuR* genes

The *Eco*RI fragment from pAR8877, harbouring *cbiW*, *-H*₆₀, *-X*, *-J*, *-C*, *D*, *-ET*, *-L*, *-F*, *-G* and *-A*, was isolated and cloned into the *Eco*RI site of pKK223.3. The resultant plasmid, pAR8980, therefore contained the whole *B. megaterium cob* operon except *cysG*^A, *cbiY* and *btuR* (Table 1).

Detection of cobyrinic acid-synthesizing colonies by bioassay

Bioassay plates were prepared as described previously [8] with the indicator strain (*S. typhimurium metE cysG*, AR3612). This strain is able to detect cobyrinic acid and later intermediates on the cobalamin biosynthetic pathway. Intermediates preceding cobyrinic acid do not seem to be transported into the reporter strain. The bioassay plates were prepared in minimum medium containing cysteine (50 µg/ml).

RESULTS

Complementation of *S. typhimurium cob* mutant strains by *B. megaterium cbi* genes for cobinamide production

To confirm that the *B. megaterium cbi* genes, which have been assigned functionality on the basis of sequence similarity, are essential for cobalamin biosynthesis, they were used to complement a number of defined *S. typhimurium cob* mutant strains [8] (Table 2). Plasmid pAR8754 (which harbours *cbiX*, *-J*, *-C*, *-D*, and *-ET*) complemented *S. typhimurium cbiJ* (AR3632 and AR3618) and *cbiD* (AR3728) mutant strains. However, although pAR8754 complemented the *cbiD* strain, complementation was achieved only after 48 h of growth rather than the normal 24 h, whereas another *cbiD* strain, AR3718, was complemented neither by pAR8754 nor by plasmid pER180, which carries only the *B. megaterium cbiD* downstream of a *tac* promoter.

The *S. typhimurium cysG* strain AR3612, which is deficient only in its transmethylese activity [8], in other words the strain can be thought of as *cysG*^A-negative and *cysG*^B-positive, was complemented by plasmid pAR8766 carrying the large *B. megaterium cob* locus, as well as by pAR8799 harbouring only the putative *B. megaterium cysG*^A (Table 2). The *S. typhimurium cbiL* strain AR3717 was complemented by plasmid pAR8766 as well as by pER110, which harbours only *cbiL*. The *S. typhimurium cbiF* strain (AR2720) was also complemented by pAR8766. However, the *S. typhimurium cbiA* mutant strain (AR2719) was not complemented by pAR8766 but was complemented by pER187, which carries only *cbiA* under the control of the *tac* promoter. Thus it seems that *cbiA* is not transcribed/translated from pAR8766. No *S. typhimurium cbiC*, *-G* or *-H* mutants were available. These results demonstrate that the *B. megaterium cob* operon contains functional *cbiE*, *-J*, *-D*, *-L*, *-F* and *-A* and *cysG*^A genes.

Synthesis of cobyrinic acid in *E. coli* by expression of the *B. megaterium cob* operon

It has been shown in a previous study that *E. coli* strains harbouring the appropriate *S. typhimurium cobI* genes (*cbiA*, *-C*, *-D*, *-E*, *-T*, *-F*, *-G*, *-H*, *-J*, *-K*, *-L* and *-P*) confer upon the host the ability to synthesize high concentrations of cobyrinic acid [8], an

Table 2 Complementation of *S. typhimurium cob* mutants with *B. megaterium cbi* genes

Strain pAR8766 is *B. megaterium cbiW-H*₆₀-*X-J-C-D-ET-L-F-G-A-cysG*^A-*cbiY-btuR*-><-ORF1-ORF2; strain pAR8754 is *S. typhimurium cbiA* and *B. megaterium cbiX-J-C-D-ET*

<i>S. typhimurium</i> mutant strain	Complementation of <i>S. typhimurium</i> mutant strains after transformation with		
	pAR8766	pAR8754	Others
AR3612 (<i>cysG metE</i>)	Y		pAR8799 (<i>cysG</i> ^A): Y
AR3717 (<i>cbiL</i>)	Y	N	pER110 (<i>cbiL</i>): Y
AR2720 (<i>cbiF</i>)	Y	N	
AR3620 (<i>cbiF</i>)		N	
AR3618 (<i>cbiJ</i>)		Y	
AR3632 (<i>cbiJ</i>)		Y	
AR2719 (<i>cbiA</i>)	N		pER187 (<i>cbiA</i>): Y
AR3718 (<i>cbiD</i>)		N	pER180 (<i>cbiD</i>): N
AR3728 (<i>cbiD</i>)		Y (weak)	
AR2709 (<i>cbiP</i>)		N	
AR2206 (<i>cbiB</i>)		N	
AR3628 (<i>cobS</i>)		N

Figure 1 Diagrammatic representation of some of the constructs used to investigate cobalamin biosynthesis from the *B. megaterium* *cobI* operon

A large X in the construct represents where a mutation has been introduced into the operon. The right-angled arrow represents the location of the *tac* promoter. The amount of cobalamin that these constructs promote when transformed into the appropriate *E. coli* strain in the presence of IPTG is also shown. Abbreviation: OD, attenuation unit.

Table 3 Cobyric acid production from the *B. megaterium* *cob* operon

The cultures were performed under anaerobic conditions in M9 medium/chloramphenicol/ampicillin/1 mg/l $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ /0.1 g/l yeast extract. The experiments were repeated several times and an average value is given for each. All strains contained the *S. typhimurium* *cbiP* gene. The difference between strains ER241 and AR8937 is the *tac* promoter added in 5' of the *cob* operon in the latter. Abbreviation: n.d., not detected.

Strain	Genes of the <i>B. megaterium</i> <i>cob</i> operon mutated or deleted	Gene added back	Cobyric acid production (pmol per attenuation unit)	
			In uninduced conditions	In induced conditions
ER241			550	600
AR8937			400	660
ER164	<i>cbiD</i>		n.d.	n.d.
ER167	<i>cbiD</i>	<i>cbiD</i>	40	320
ER165	<i>cbiG</i>		0.5	0.3
ER169	<i>cbiG</i>	<i>cbiG</i>	290	680
AR8930	<i>cbiH₆₀</i> (second part)		220	400
ER144	<i>cbiW</i>		45	500
AR8925	<i>cbiX</i>		0.1	2
AR8967	<i>cbiX</i>	<i>cbiX</i>	110	440
ER176		<i>P.d cobA</i>	300	460
ER177	<i>cysG^A-cbiY-btuR</i>	<i>P.d cobA</i>	0.4	390

ability that *E. coli* by itself is inherently incapable of because it does not possess the genes for the CobI pathway (precoirrin-2 to cobinamide) [12]. By the same rationale as that used for the expression of the *S. typhimurium* CobI enzymes within *E. coli*, the *B. megaterium* *cob* operon was placed under the control of a *tac* promoter by cloning the 16 kb *cob* fragment into pKK223-3, a procedure that yielded plasmid pAR8877 (Table 1 and Figure 1). However, as the *B. megaterium* 16 kb fragment does not contain *cbiP*, this plasmid would not be expected to promote the synthesis of cobalamin, the earliest intermediate that can be detected by bioassay, after transformation into competent *E. coli*. Thus the *S. typhimurium* *cbiP* was introduced on a second plasmid named pAR8849; this plasmid also contained the repressor *lacI^r* of the *tac* promoter. The strain containing the two plasmids pAR8877 and pAR8849 (AR8937) was grown under anaerobic conditions and was found to produce cobalamin at 400 pmol per attenuation unit when uninduced, and at 660 pmol per attenuation unit when induced with IPTG (Figure 1 and Table 3). When the same strain was grown aerobically, no

cobyric acid could be detected. Thus cobalamin synthesis with the *B. megaterium* *cobI* operon in *E. coli*, like the synthesis of cobalamin in *S. typhimurium*, is dependent on anaerobicity.

Region upstream of *cbiW* contains a *cob* promoter

To investigate the role of the region upstream of *cbiW* on expression in *E. coli*, strain ER241 was constructed. This strain contains both pAR8849 (*lacI^r* and *cbiP* genes) and pAR8766 (*B. megaterium* *cobI* genes but no *tac* promoter) (Figure 1) and, under anaerobic growth conditions, is able to produce cobalamin at approx. 500–600 pmol per attenuation unit in the presence or absence of IPTG (Figure 1 and Table 3). This demonstrates that the *B. megaterium* *cob* genes are efficiently transcribed in the given culture conditions from their original promoter(s) situated upstream of *cbiW*. When the *B. megaterium* *cob* genes were cloned under the control of the *tac* promoter (pAR8877) (Figure 1), as in strain AR8937, significant levels of cobalamin were synthesized under both induced (660 pmol per attenuation unit)

and non-induced conditions (400 pmol per attenuation unit) (Table 3). However, deletion of *cbiW* and its upstream region from the *tac*-regulated plasmid (producing pAR8846) resulted in a large decrease in cobyrinic acid synthesis in uninduced *E. coli* (ER144) (Table 3). The level of cobyrinic acid synthesis was restored by the addition of IPTG to the medium. This demonstrates that the region upstream of *cbiW* promotes constitutive expression of the *B. megaterium* *cob* operon.

***B. megaterium* *cbiD* and *cbiG* are necessary for cobyrinic acid production**

Although complementation studies on the *S. typhimurium* *cob* mutants had defined essential roles for the *B. megaterium* *cbiE*, *-J*, *-L*, *-F* and *-A* and *cysG*, the result for *cbiD* was not convincing. Furthermore no defined mutants for *cbiC*, *-G*, *-H*, *-W*, *-X* or *-Y* were available to demonstrate the need for these genes in cobalamin biosynthesis. To circumvent this problem, defined *cobI* mutants were generated within the recombinant cobyrinic acid-producing *E. coli* strain AR8937 by genetic modification of the main *cobI*-harbouring plasmid pAR8877. Previously it had been demonstrated that, in the *S. typhimurium* cobalamin biosynthetic pathway, the presence of both *cbiD* and *cbiG* in a functional form was essential for cobyrinic acid production [8]. To determine whether the putative *B. megaterium* *cbiD* and *cbiG* genes are necessary for cobyrinic acid biosynthesis in the *B. megaterium* *cob* operon, mutations were introduced separately into *cbiD* and *cbiG* within pAR8877 (Figure 1) (see the Materials and methods section). The derivative of pAR8877 with a mutation in *cbiD* was called pAR8909 and when this was introduced into *E. coli*, to give strain ER164, no detectable amounts of cobyrinic acid were produced (Figure 1 and Table 3). When strain ER164 was transformed by an additional plasmid carrying a wild-type *B. megaterium* *cbiD* (pER156), the resulting strain (ER167) produced cobyrinic acid at 40 pmol per attenuation unit constitutively, whereas 320 pmol per attenuation unit was produced when the strain was induced with IPTG (Table 3). The derivative of pAR8877 that contained the mutated *cbiG* was called pAR8921 (Figure 1) and when transformed into *E. coli* gave strain ER165, which produced barely detectable levels of cobyrinic acid (less than 1 pmol per attenuation unit) in the presence or absence of IPTG (Table 3). However, when the wild-type *B. megaterium* *cbiG* was reintroduced back into the strain on plasmid pER160, cobyrinic acid production was restored to normal levels (Table 3). These results demonstrate that the putative *B. megaterium* *cbiD* and *cbiG* genes are necessary for cobyrinic acid production in the *B. megaterium* *cob* operon.

***B. megaterium* *cbiX* is necessary for cobyrinic acid production**

To determine whether *B. megaterium* *cbiX*, to which no function has been assigned, is necessary for cobyrinic acid production, a mutation was introduced into *cbiX* in a similar manner to that previously described for both *cbiD* and *-G*. The derivative of pAR8877 harbouring the mutated *cbiX* was called pAR8911 (Figure 1) and when this plasmid was transformed into *E. coli* it gave strain AR8925, which produced cobyrinic acid at only 2 pmol per attenuation unit (Table 3). When a copy of the wild-type *cbiX*, on pAR8962, was reintroduced back into the strain (giving AR8967), the level of cobyrinic acid produced under inducing conditions was 440 pmol per attenuation unit (Table 3). This established *cbiX* as an essential gene for cobalamin biosynthesis within the *cob* operon.

cbiW*, *cbiY* and the 3' extension of *cbiH*₆₀ are not required for cobalamin biosynthesis in *E. coli

The removal of the first 1914 bp of the *B. megaterium* 16 kb DNA fragment within pAR8877 gave rise to pAR8846. The deleted region spanned the whole region situated upstream of the first gene of the *cob* operon, *cbiW*, as well as the first 50 codons of the 127 codon *cbiW* itself (Figure 1). When pAR8846 was transformed into *E. coli* together with *cbiP* it yielded strain ER144 but the deletion of the 1914 bp had little effect on cobyrinic acid synthesis as the strain produced 500 pmol per attenuation unit when grown anaerobically in the presence of IPTG (Figure 1 and Table 3). As described above, the low level of cobyrinic acid observed when ER144 was grown in the absence of IPTG (45 pmol per attenuation unit) was probably a consequence of the deletion of the *cob* operon promoter.

A deletion of the distal (3') end of the *cob* operon, which includes *cysG*^A-*cbiY*-*btuR*, resulted in the construction of pAR8980 (Figure 1). When this plasmid was co-transformed into *E. coli* with a compatible plasmid harbouring *P. denitrificans* *cobA* (the equivalent of the *cysG*^A portion of the *E. coli* *cysG* gene; pER146) strain ER177 was generated. No major difference in cobyrinic acid production was observed between this strain and the strain containing the complete *B. megaterium* *cob* operon and the *P. denitrificans* *cobA* (ER176), as both strains produced cobyrinic acid at approx. 400 pmol per attenuation unit when they were grown anaerobically in the presence of IPTG (Table 3). When *cbiH*₆₀ was truncated within the *cob* operon by mutagenesis such that the mutated *cbiH*₆₀ encoded a truncated protein lacking the last 162 amino acid residues but leaving the main CbiH domain intact (pAR8916) (Figure 1), the mutation did not decrease cobyrinic acid synthesis significantly because the strain harbouring mutant *cbiH*₆₀ (AR8930) still produced 400 pmol per attenuation unit when grown anaerobically in the presence of IPTG (Table 3).

DISCUSSION

The functional complementation of defined *S. typhimurium* *cob* mutants and genetically engineered *E. coli* strains capable of cobyrinic acid synthesis has confirmed the role of *B. megaterium* *cbiA*, *-D*, *-ET*, *-F*, *-G*, *-H*, *-J*, *-L* and *-F* and *cysG*^A in cobalamin biosynthesis. However, *B. megaterium* *cbiD* did not complement one of the two *S. typhimurium* *cbiD* strains and only complemented the other *cbiD* strain very slowly. This was surprising as the *B. megaterium* and *S. typhimurium* CbiD protein sequences exhibited a reasonable degree of similarity (initn = 614, initl = 318; 34.4% identity) (the preceding paper [11]). It is possible that each of the two plasmids carrying the *B. megaterium* *cbiD* gene that were used in the complementation experiments in *S. typhimurium* did not express adequate amounts of CbiD, or that the protein did not fold properly within the cell. However, this is unlikely because cobyrinic acid synthesis from the *B. megaterium* *cob* operon in *E. coli* was dependent on the presence of *cbiD*. Alternatively, the function of the *B. megaterium* CbiD might require the presence of a gene product such as the C-terminus of CbiH₆₀, which is not found in *S. typhimurium*.

Within the *B. megaterium* *cob* operon, *cbiX* was shown to be an essential component of the cobalamin biosynthetic apparatus because inactivation of the gene resulted in a decrease in cobalamin biosynthesis to 1/200. Although *cbiX* has no sequence similarity to any of the the *S. typhimurium* *cbi* genes, the presence of a polyhistidine motif within the protein (the preceding paper [11]) and its recently described ability to complement an *E. coli* *cysG*^B strain [15] suggest that it is functionally equivalent to *cbiK*, the putative precorrin-2 cobalt-chelatase.

Deletion of *cbiW* and *-Y* and the C-terminal portion of CbiH₆₀ (deletion of the terminal 162 residues) had no detrimental effect on corrin biosynthesis in *E. coli*. Although such deletion experiments also suggest the presence of a putative *B. megaterium* cobalamin promoter upstream of *cbiW*, the genes themselves are not essential for cobalamin biosynthesis. They might be involved in regulation of the pathway or perhaps required in the reduction of the central cobalt ion during ring contraction. However, if it is the latter, *E. coli* must possess an equivalent protein as the engineered *E. coli* strains are able to synthesize cobyrinic acid in their absence.

It has been convincingly argued that *E. coli* lost its Cbi pathway during its evolution, although it has maintained the *cobII* and *cobIII* operons [12]. The instigation of corrin biosynthesis within *E. coli* by transformation with a *cobI* operon from either *B. megaterium* or *S. typhimurium* can therefore be looked on as restoration of a previously lost ability. However, what is particularly striking about the restoration of corrin synthesis by the *B. megaterium cobI* operon is that it occurs only when the genetically modified *E. coli* are grown anaerobically. Two reasons can be put forward to explain why a cobalamin biosynthetic pathway from an obligate aerobe functions only under anaerobic conditions in a facultative organism. First, when *E. coli* is grown aerobically the oxidative environment might prove to be detrimental to the highly labile early corrin intermediates. In *B. megaterium* the labile intermediates might be protected by other ancillary protein factors. Alternatively, *E. coli* might encode a protein, expressed only under strictly anaerobic conditions, that might be involved in the oxidation–reduction of the central cobalt ion, a process thought to mediate the ring-contraction process [7].

The identification of the functional genes within the *B. megaterium cob* operon together with the equivalent identification of the *S. typhimurium cbi* genes has laid an important foundation from which anaerobic cobalamin biosynthesis can be completely

elucidated. Previously, the biosynthesis of vitamin B₁₂ has been described as the Mount Everest of biosynthetic problems [6]. However, as with all mountains, there is often more than one route to the summit and, as might be expected, the route to the top of such a large mountain is proving to be more challenging in the absence of oxygen.

REFERENCES

- 1 Roth, J. R., Lawrence, J. G., Rubenfield, M., Kieffer-Higgins, S. and Church, G. M. (1993) *J. Bacteriol.* **175**, 3303–3316
- 2 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
- 3 Debussche, L., Thibaut, D., Cameron, B., Crouzet, J. and Blanche, F. (1993) *J. Bacteriol.* **175**, 7430–7440
- 4 Blanche, F., Thibaut, D., Debussche, L., Hertle, R., Zipfel, F. and Müller, G. (1993) *Angew. Chem. Int. Ed. Engl.* **32**, 1651–1653
- 5 Warren, M. J. and Scott, A. I. (1990) *Trends Biochem. Sci.* **15**, 486–491
- 6 Battersby, A. R. (1994) *Science* **264**, 1551–1557
- 7 Scott, A. I. (1994) *Heterocycles* **39**, 471–476
- 8 Raux, E., Lanois, A., Levillayer, F., Warren, M. J., Brody, E., Rambach, A. and Thermes, C. (1996) *J. Bacteriol.* **178**, 753–767
- 9 Raux, E., Thermes, C., Heathcote, P., Rambach, A. and Warren, M. J. (1997) *J. Bacteriol.* **179**, 3202–3212
- 10 Scott, A. I., Stolowich, N. J., Wang, J. J., Gawatz, O., Fridrich, E. and Muller, G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14316–14319
- 11 Raux, E., Lanois, A., Warren, M. J., Rambach, A. and Thermes, C. (1998) *Biochem. J.* **335**, 159–166
- 12 Lawrence, J. G. and Roth, J. R. (1996) *Genetics* **142**, 11–24
- 13 Bonnett, R., Godfrey, J. M. and Redman, D. G. (1969) *J. Chem. Soc. C*, 1163–1166
- 14 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- 15 Beck, R., Raux, E., Thermes, C., Rambach, A. and Warren, M. J. (1997) *Biochem. Soc. Trans.* **25**, 77S
- 16 Roessner, C. A., Spencer, J. B., Ozaki, S., Min, C., Atshaves, B. P., Nayar, P., Anousis, N., Stolowich, N. J., Holderman, M. T. and Scott, A. I. (1995) *Protein Express. Purif.* **6**, 155–163